THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Advancing CRISPR technologies to engineer yeast metabolism

RAPHAEL FERREIRA

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Biology and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Advancing CRISPR technologies to engineer yeast metabolism

Raphael Ferreira
Gothenburg, Sweden 2019
ISBN 978-91-7905-148-8

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4615
ISSN 0346-718X

Division of Systems and Synthetic Biology
Department of Biology and Biological Engineering
Chalmers University of Technology
SE-412 96
Gothenburg, Sweden
Telephone: + 46 (0) 31 772 10 00

Cover: Synthwave representation of CRISPR technology and yeast metabolism with Gothenburg in the background.

Printed by Chalmers Reproservice
Gothenburg, Sweden 2019

Advancing CRISPR technologies to engineer yeast metabolism

Raphael Ferreira
Department of Biology and Biological Engineering
Chalmers University of Technology

Abstract

The advent of genetic engineering tools has initiated an era of manipulating microorganisms for the production of valuable compounds for our society. Precise engineering of these microbes commonly requires introducing genetic modifications such as gene deletion, overexpression, and accurate regulation in order to enhance the production of the compound of interest. In this context, the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 technology, adapted from the prokaryotic adaptive immune system, has revolutionized our ability to manipulate a broad range of living organisms. In contrast to other methods, this technology works like a molecular pair of scissors (Cas9) which is guided by a programmable RNA (gRNA) molecule binding at a specific location in the DNA. The programmability and time-efficiency offered by this technology have in the recent years been successfully exploited in rewiring the metabolic network to enhance the production of metabolites used in various areas of industrial biotechnology.

In this thesis, I present several studies applying the technological diversity provided by CRISPR in the context of building efficient yeast cell factories for the production of oleochemicals -sustainable substitutes for plant derived lipids. Since oleochemicals derive from lipid products, the main engineering strategies presented essentially focus on fatty acid metabolism and its precursors. First, we exploited CRISPR/Cas9 endonuclease capacity to extensively remodel yeast lipid metabolism. We showed that the disruption of several metabolic fluxes allows to overcome the main limiting steps in fatty acid biosynthesis and favors the production of free fatty acids and triacylglycerols, two important precursors for the production of oleochemicals. Second, we harnessed the ability to precisely regulate genes using the catalytically deactivated form of the Cas9 protein (dCas9) coupled to transcription factors for fine-tuning the expression of genes involved in lipid biogenesis. Additionally, we proposed a framework for dCas9-based applications based on computational techniques for predicting key genes potentially favoring the production of yeast endogenous metabolites. Finally, we expanded the CRISPR repertoire by building new tools to accelerate yeast cell factory design. We exploited a Type I CRISPR-associated endoribonuclease for multiplex genome engineering and transcriptional regulation *via* processing an RNA transcript into multiple gRNAs, and we developed a computational tool for designing gRNAs targeting multiple loci at once. In summary, the work presented in this thesis provides various ways to efficiently engineer yeast metabolism by exploiting the diversity of CRISPR technologies, as well as new tools to the community for future engineering strategies.

Keywords: CRISPR, metabolic engineering, *Saccharomyces cerevisiae*, oleochemicals

List of publications

This thesis is based on the work contained in the following papers and manuscripts:

I. **Ferreira, R.,** David, F. and Nielsen, J., 2018. Advancing biotechnology with CRISPR/Cas9: recent applications and patent landscape. *Journal of Industrial Microbiology & Biotechnology*, pp.1-14.

II. **Ferreira, R.*,** Teixeira, P.G.*, Siewers, V. and Nielsen, J., 2018. Redirection of lipid flux toward phospholipids in yeast increases fatty acid turnover and secretion. *Proceedings of the National Academy of Sciences*, 115(6), pp.1262-1267.

III. **Ferreira, R.,** Teixeira, P.G., Gossing, M., David, F., Siewers, V. and Nielsen, J., 2018. Metabolic engineering of *Saccharomyces cerevisiae* for overproduction of triacylglycerols. *Metabolic Engineering Communications*, 6, pp.22-27.

IV. Jensen, E.D.*, **Ferreira, R.***, Jakočiūnas, T., Arsovska, D., Zhang, J., Ding, L., Smith, J.D., David, F., Nielsen, J., Jensen, M.K. and Keasling, J.D., 2017. Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. *Microbial Cell Factories*, 16(1), p.46.

V. **Ferreira, R.,** Skrekas, C., Hedin, A., Sánchez, BJ., Nielsen, J., and David, F. Model-assisted fine-tuning of central carbon metabolism in *Saccharomyces cerevisiae* through dCas9-based regulation **[under review]**

VI. **Ferreira, R.,** Skrekas, C., Nielsen, J. And David, F., 2017. Multiplexed CRISPR/Cas9 genome editing and gene regulation using Csy4 in *Saccharomyces cerevisiae*. *ACS Synthetic Biology*, 7(1), pp.10-15.

VII. **Ferreira, R.*,** Gatto, F.* and Nielsen, J., 2017. Exploiting off-targeting in guide-RNAs for CRISPR systems for simultaneous editing of multiple genes. *FEBS Letters*, 591(20), pp.3288-3295.

*** Contributed equally**

Additional publications not included in the thesis:

VIII. Smith, C.J.*, Castanon, O.*, Said, K., Volf, V., Khoshakhlagh, P., Hornick, A., **Ferreira, R.**, Wu, C.T., Güell, M., Garg, S., Myllykallio, H. and Church, GM (2019) Enabling large-scale genome editing by reducing DNA nicking. bioRxiv, p.574020.

IX. Lau, A.N., Li, Z., Danai, L.V., Westermark, A., Darnell, A.M., **Ferreira, R.**, … Yilmaz, O.H., Vander Heiden, M.G., Dissecting metabolic heterogeneity in pancreatic ductal adenocarcinoma **[Manuscript]**

X. **Ferreira, R.*,** Hedin, A.*, Jens Nielsen, and David, F., Biosensor-Assisted Laboratory Evolution of Malonyl-CoA production in *Saccharomyces cerevisiae* **[Manuscript]**

XI. **Ferreira, R.,** Limeta, A. and Nielsen, J., 2018. Tackling Cancer with Yeast-Based Technologies. *Trends in Biotechnology*

XII. Robinson, J.*, Kocabaş, P.*, Wang, H.*, Cholley, P.E.*, …, **Ferreira, R.**, Domenzain, I., Svensson, T., Palsson, B.O., Mardinoglu, A., Hansson, L., Uhlen, M., Nielsen J., An Atlas of Human Metabolism **[Under review]**

XIII. Teixeira, P.G., **Ferreira, R.**, Zhou, Y.J., Siewers, V. and Nielsen, J., 2017. Dynamic regulation of fatty acid pools for improved production of fatty alcohols in *Saccharomyces cerevisiae*. *Microbial Cell Factories*, 16(1), p.45.

XIV. Tippmann, S., **Ferreira, R.**, Siewers, V., Nielsen, J. and Chen, Y., 2017. Effects of acetoacetyl-CoA synthase expression on production of farnesene in *Saccharomyces cerevisiae*. *Journal of Industrial Microbiology & Biotechnology*, 44(6), pp.911-922.

XV. Gatto, F., **Ferreira, R.,** and Nielsen, J., Pan-cancer analysis of the metabolic reaction network **[Under review]**

XVI. Bergenholm, D.*, Dabirian, Y.*, **Ferreira, R.***, and Nielsen, J., Rational gRNA design based on transcription factor binding data **[Manuscript]**

XVII. Castanon, O.*, Smith, C.J.*, Khoshakhlagh, P., **Ferreira, R.**, …, Myllykallio, H and Church, GM CRISPR-mediated biocontainment technologies **[Manuscript]**

XVIII. Domenzain, I., Sánchez, B.J., Anton, M., **Ferreira, R.**, Kerkhoven, E.J., and Nielsen, J. GECKO 2.0: A generalized toolbox for enhancing genome-scale metabolic models with enzymes constraints. **[Manuscript]**

XIX. Skrekas, C., **Ferreira, R.**, and David, F. Fluorescence Activated Cell Sorting as a tool for recombinant strain screening **[Manuscript]**

Contribution summary

I. Conceptualized the review, carried out the literature search and wrote most of the manuscript.

II. Shared with the co-first author: Conceptualization of the study, carried out the experiments, analyzed the data, and wrote the manuscript.

III. Conceptualized the study, carried out the experiments, analyzed the data, and wrote the manuscript.

IV. Participated in the conceptualization of the study, carried out the experiments on the inducible system and triacylglycerol quantification, and analyzed the data.

V. Conceptualized the study, carried out parts of the experiments, analyzed the data, and wrote the manuscript.

VI. Conceptualized the study, carried out parts of the experiments, and wrote the manuscript.

VII. Conceptualized the study, carried out the wet lab experiments, and wrote the manuscript together with the co-first author.

——

VIII. Performed parts of the data analysis and helped drafting the manuscript.

IX. Performed parts of the experimental work.

X. Conceptualized the study and wrote parts of the manuscript.

XI. Conceptualized the review, carried out the literature search and wrote parts of the paper

XII. Worked on the enzyme constrained part of the study and wrote parts of the manuscript.

XIII. Performed parts of the experimental work.

XIV. Performed parts of the experimental work.

XV. Performed parts of the *in silico* work.

XVI. Participated in the conceptualization of the study, carried out parts the experiments and wrote parts of the manuscript.

XVII. Performed parts of the data analysis.

XVIII. Helped with part of the *in silico* work related to the application to human models.

XIX. Reviewed the manuscript and helped with the figures.

Preface

This dissertation serves as partial fulfillment of the requirements to obtain the degree of Doctor of Philosophy at the Department of Biology and Biological Engineering at Chalmers University of Technology. The PhD studies were carried out between March 2015 and August 2019 at the Division of Systems and Synthetic Biology (SysBio) under the supervision of Jens Nielsen, and co-supervised by Verena Siewers and Florian David. The thesis was examined by Christer Larsson. It was mainly funded by the Swedish Foundation for Strategic Research.

Raphael Ferreira
July 2019

Table of content

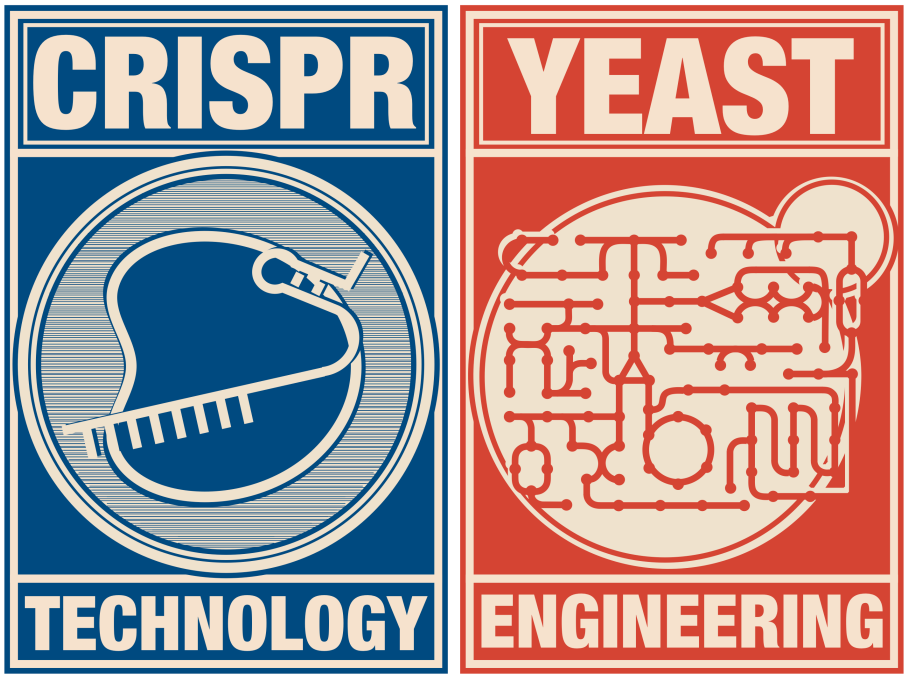
Abbreviations

aTc	anhydrotetracycline
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Cas9	CRISPR associated protein 9
dCas9	Dead-Cas9
VPR	VP64-p65-Rta
gRNA	Guide-RNA
tracrRNA	Trans-activating RNA
crRNA	CRISPR RNA
PAM	Protospacer Adjacent Motif
ER	Endoplasmic Reticulum
PA	Phosphatidic acid
PPP	Pentose Phosphate Pathway
PS	Phosphatidylserine
CL	Cardiolipin
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
PI	Phosphatidylinositol
DBTL	Design, Build, Test, and Learn
FFA	Free Fatty Acid
TAG	Triacylglycerol
SE	Sterol Ester
FACS	Fluorescence-activated cell sorting
NGS	Next-generation sequencing
GEM	Genome-scale metabolic model
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
GEM	Genome scale metabolic models
FA	Fatty acids

La lutte elle-même vers les sommets suffit à remplir un cœur d'homme. Il faut imaginer Sisyphe heureux.

Albert Camus, Le mythe de Sisyphe

Chapter I: Engineering yeast metabolism



1.1 Prologue

The concept of microbial biotechnology, when defined as the manipulation of microorganisms to convert raw materials into valuable products, can be traced back to the early ages of human society (1, 2). Such utilization exclusively relied on basic fermentation principles from a narrowed number of organisms, with baker's yeast figuring prominently as a result of its fermentative capability to yield beer, bread, and wine. Arguably, with the polyvalence of its fermented products, domestication of baker's yeast can be praised as one of the most critical contributors to the societal growth: with beer and wine as necessary means of diplomacy, recreation and socialization, and bread as a universal nutritious food, also core to several religious cultures. Though, these processes were uncharacterized until the mid-19th century, attributed to the now obsolete theory of spontaneous generation (Fig. I-1). Among other seminal works that provided better understanding of the biology behind fermentation, Louis Pasteur's work on deciphering yeast as the catalyst behind the alcoholic fermentation has initiated the discipline of microbiology (3). The advent of this discipline provided the fundamental characterization of microorganisms on the subject of biochemistry, physiology, and cell biology, as well as enhanced the search for new species with attracting properties. Consequently, the war economy, and more specifically the manufacturing of explosives, greatly benefited from microbial production of acetone and glycerol by *Clostridium acetobutylicum* and *Saccharomyces cerevisiae*, respectively (4, 5). In the medical field, Alexander Fleming's serendipitous discovery of penicillin, an

antibacterial compound secreted by *Penicillium chrysogenum*, is a landmark in the history of treating bacterial diseases (6). While these examples highlight the exploitation of intrinsic microbial qualities shaped through evolution to human advantages, these microbes naturally only produce a small amount of these compounds. Thus, scientists' past efforts to improve the production have been essentially focused on altering the physiological state of these organisms by tweaking their cultivation methods, e.g. addition of sulfite in the media to inhibit the conversion of acetaldehyde to ethanol to rewire fluxes towards glycerol (7) (Fig. I-1).

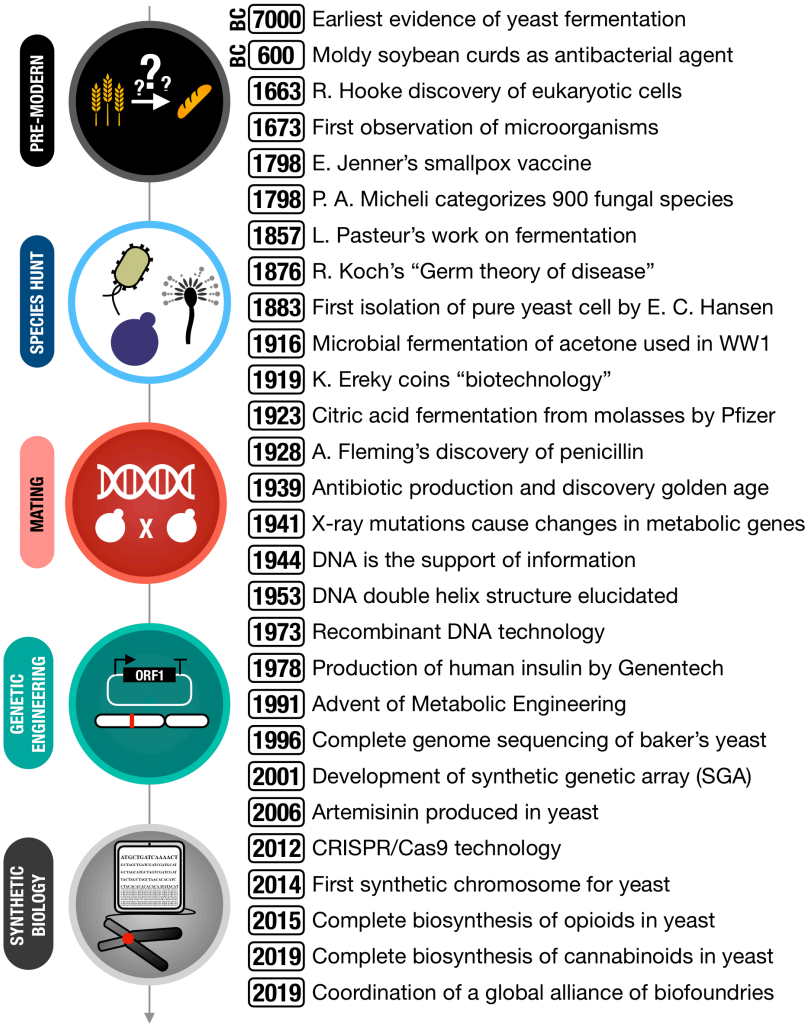


Figure I-1. Timeline of key event in microbial biotechnology. References: **(1, 3, 6, 8–28)**

The twentieth century also marks important advances in our fundamental understanding of genetics via radiation technologies, e.g. the elucidation of the protein structure (29) and the double helix structure of deoxyribonucleic acid (DNA) (8). Additionally, these technologies enabled the characterization of a myriad of genetic mapping and physiologic characterizations through their mutagenesis properties, e.g. elucidation of genes encoding for proteins catalyzing different sugar consuming reactions (30–33) (Fig. I-1). Rather than focusing on culture conditions, exposing microorganisms to random

mutagenesis allowed targeted selection of strains, with genetic variants leading to desired physiological properties.

A major technological revolution occurred in 1973 when Stanley Cohen co-jointly with Herbert Boyer developed a technique for merging together strands of DNA and artificially introduce them into an organism, coined recombinant DNA technology (9). This method allowed for the first time the expression of heterologous genes into a given host. Consequently, the first genetic engineering company, Genentech Inc. founded by Herbert Boyer and Robert Swanson, famously established the first production of the synthetic human protein (somatostatin) in *Escherichia coli*, with identical properties to the naturally occurring one (34). In 1987, Novo produced human insulin in *S. cerevisiae* as a substitute for their insulin obtained from animal origin (35). These engineering advances also enabled better control over metabolic pathways in order to enhance the production of natural and new compounds. This consequently extended the focus towards model organisms such as *E. coli* and *S. cerevisiae*, for the production of relevant compound rather than natural producers.

Strain optimization using synthetic biology and systems biology

The advent of genetic engineering tools along with modern information processing technologies, have initiated an era of precise control of microbial fermentation for the production of diverse compounds. Here, these microbial producers, coined cell factories, are not only engineered to express relevant compounds but also extensively characterized through incorporated approaches aiming to understand how to further enhance production processes. Thus, in the beginning of the 1990s, a discipline studying the directed enhancement of metabolic pathways to achieve a specific function, e.g. high production of a desired product, was initiated (coined metabolic engineering) (20, 36). Its principles lie on the rational application of the knowledge obtained from fundamental biology combined with the efficiency of molecular biology tools in cell factory engineering.

Distinctively, sequential and targeted optimization of strain engineering is iteratively applied through the so-called **Design**, **Build**, **Test**, and **Learn** (DBTL) cycle (Figure 1) (37). In the **design** step, the desired target compound defines which chassis to select as well as its cellular functions to engineer. For example, based on the complexity of the target product, one might start from an organism originally producing the compound or from a model organism easier to engineer. The systematic analysis of vast biological datasets can enable the prediction of biological entities' behavior upon metabolic rerouting and majorly contributes to the accuracy of strain **design**. For example, genome-scale metabolic models (GEMs), which are genome-wide mathematical models, are attractive tools for predicting key reactions that potentially favor the production of a metabolite of interest (38–40). Furthermore, technological progresses in mapping the various layers of cellular processes can also be integrated into strain design, e.g. genomics for genome mapping, transcriptomics for RNA mapping, or metabolomics for mapping metabolites etc.

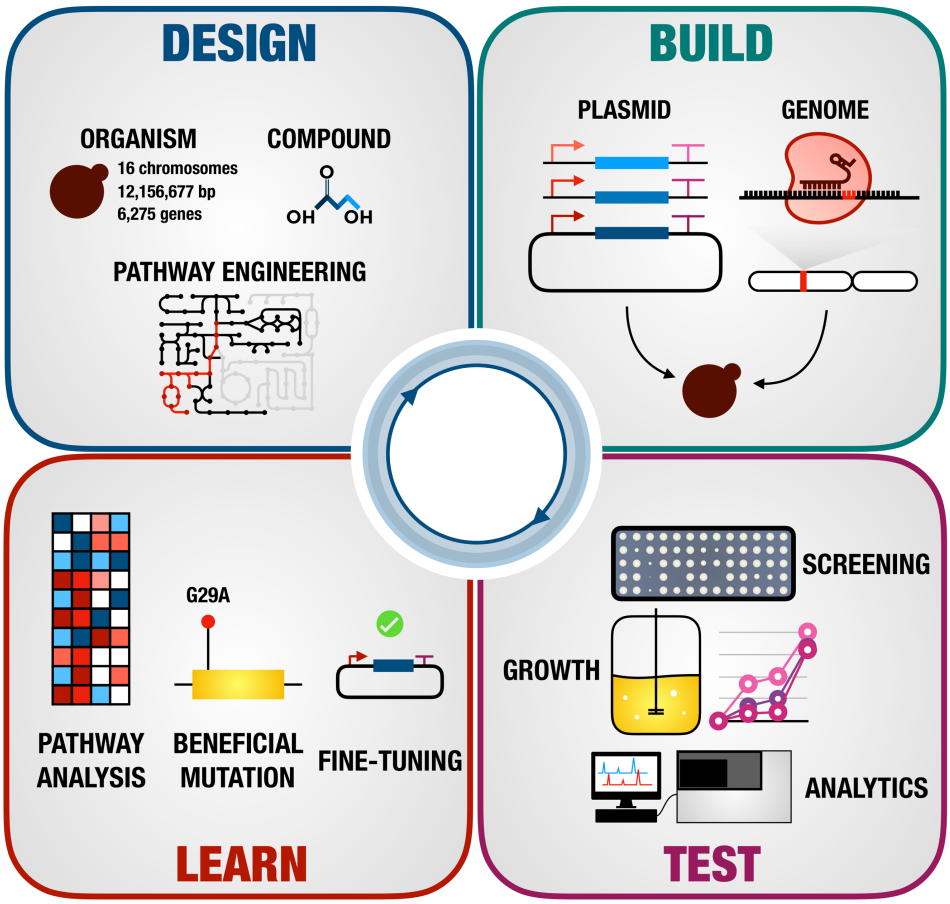


Figure I-2. The Design, Build, Test, and Learn (DBTL) cycle for cell factory enhancement.

Based on the design, appropriate biological parts are selected for the **building step**. For example, control over gene expression requires using libraries of promoters or increasing the copy of a given gene. This step generally lies on molecular biology approaches to assemble multiple DNA fragments together and transform them into the selected host organism (41, 42). Additionally, besides rewiring existing (or creating new) pathways through these building processes, tools that allow sensitive and rapid screening of candidate strains can be also integrated to the design (43, 44). For example, transcription-factor based biosensors allow to monitor the intracellular concentration of a specific compound when coupled to a fluorescent gene and can ultimately accelerate the throughput of evaluating the desired phenotype. An example of utilization of a biosensor will be described in **Chapter III**.

Next, the **test** phase assesses the performance of the engineered strains for different parameters, e.g. measuring the production of the desired product through the determination of titer, rate, and yield, using analytic techniques; specific growth behavior and cell physiology using bioreactors; or high-throughput screening of strain libraries by tracking single cell GFP fluorescence using fluorescence activated cell sorting (FACS). Finally, the most optimized strains, i.e. strains giving the best expected output of the study, are analyzed to **learn** which characteristics lie behind these superior performances as well as how to

further improve them (Fig. I-2). For example, genome or gene sequencing to identify causal mutations in these strains, or, RNA sequencing analysis to identify genes and pathways causing physiological changes in these strains (37).

1.2 The last genetic engineering technology

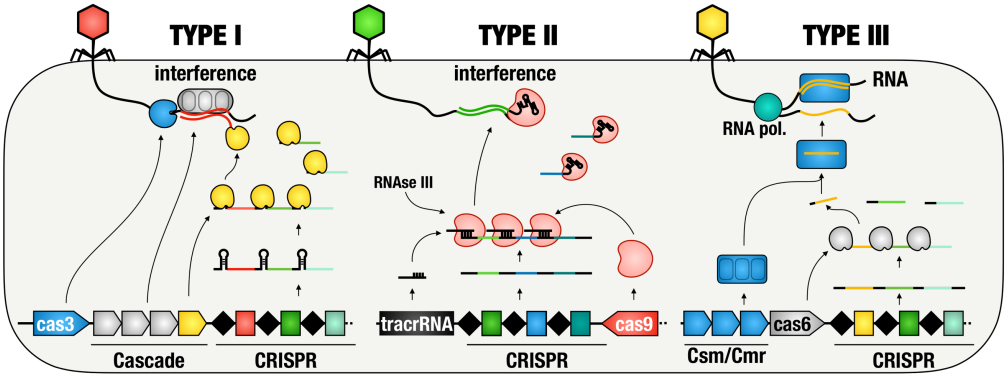
Metabolic pathway optimization towards a product of interest commonly requires deletions of multiple genes, e.g. competitive metabolic pathways, which, in yeast, is traditionally performed through iterative cycles of genetic marker integration and removal (45). However, these tools are commonly experienced with difficulties in design, synthesis, and time-efficiency which altogether prevents fast strain engineering (45).

In the last seven years, one particular technique, named Clustered Regularly Interspaced Short Palindromic Repeats (**CRISPR**), has been efficiently addressing most of the above-mentioned challenges in genetic engineering. This technology originates from prokaryotic and archaeal immune systems to neutralize phage invasions, and has now been tailored for manipulating a broad range of living organisms (24, 46, 47), including cell factories (48).

CRISPR Principles

First reported in 1987, **CRISPR** derives from arrays found in the genome of numerous bacteria and archaea (49). These arrays are composed of DNA sequences encoding non-coding CRISPR RNAs (crRNAs), comprised of short direct repeat sequences (DR) interspaced with unique ones (spacers), and clustered together with several CRISPR-associated (*cas*) genes. These two distinct elements together form the main elements of the adaptive immunity against foreign nucleic acid invasion such as phage genomes and conjugating plasmids (50, 51). The mediated immunity is divided into two stages: *(i)* the **adaptation** stage, commonly driven by *cas1* and *cas2* genes (52), whose function is to genomically integrate elements of the phage genome into the beginning of the CRISPR loci; and *(ii)* **expression** and **interference** stage, composed of diverse subsystems involved in processing of the different spacers and degradation of the invasive viral genome through *cas* nuclease effectors. Interestingly, the constant war against phages has yielded an extreme diversification of *cas* genes and genomic locus architecture and rigorous classification of the different CRISPR types and classes remains an ongoing work (53) (54) (Box 1).

Box 1. Type I, II, and III CRISPR-systems.



Type I systems use Cas6e/Cas6f/Csy4 endoribonucleases (in yellow) to generate CRISPR RNAs (crRNAs) by cleaving at the intersection of the hairpin loops in the CRISPR direct repeat. The interference stage occurs when the crRNAs bind to its viral complementary strand (protospacer), considering the presence of the PAM. Once bound, the multiprotein Cascade complex (in grey) recruits Cas3 (in blue) for DNA degradation.

Type II systems use a trans-activating RNA (tracrRNA) to generate multiple crRNAs. This system relies on a single protein, e.g. Cas9, for the interference step.

Type III systems use Cas6 to produce mature crRNAs. As with type I, type III systems require multiple Cas proteins binding to crRNAs, called Csm/Cmr complex. The type III systems can either target the viral mRNA or DNA.

Application in genome engineering and gene regulation

Scientists have since borrowed these different CRISPR-systems to perform complex gene editing (48), with the type II system adapted from *Streptococcus pyogenes* being the most widely adopted for genetic engineering and gene regulation (Box 1) (55, 56). This technology targets DNA with an endonuclease, CRISPR associated protein 9 (Cas9), which is programmed by an RNA, guide-RNA (gRNA). The gRNA is composed of two parts: *(i)* the crRNA, i.e. the transcribed spacer, usually 20 nucleotide long and complement to the target DNA (protospacer); and *(ii)* the tracrRNA, which acts as connecting scaffold with Cas9 (24, 57) (Box 1, Fig. I-3). The crystal structure of Cas9 revealed the composition of two lobes linked with a bridge helix (BH): the target recognition (REC) whose function is involved in binding between the gRNA and the DNA target, and the nuclease lobe (NUC) containing the nucleases HNH and RuvC (58) (Fig. I-3). Once the Cas9-gRNA complex is bound to the DNA target by base pairing, Cas9 nucleases domains cleave the DNA sequence three nucleotides upstream the protospacer adjacent motif (PAM, NGG-3' for Cas9), recognized by the PAM-interacting domain (PI) located in the NUC lobe (58, 59) (Fig. I-3). The PAM sequence is juxtaposed to the protospacer region and is required for DNA cleavage (Fig. I-3) (60).

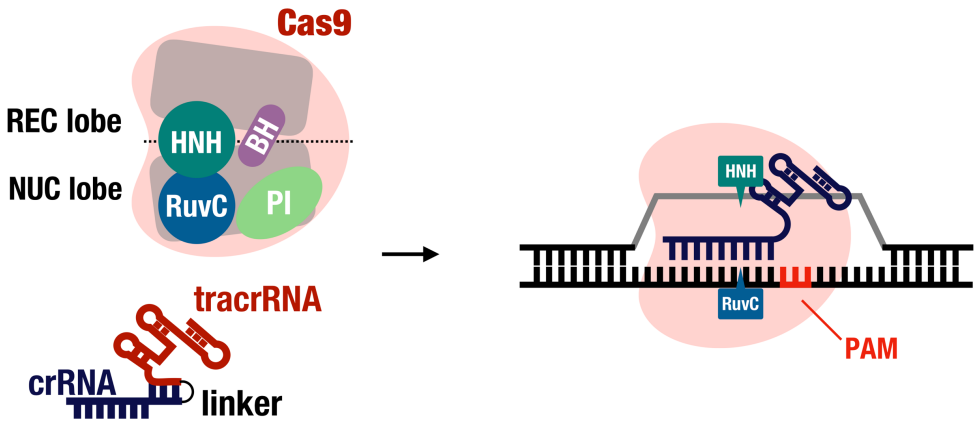


Figure I-3. The main elements of the CRISPR technology: Cas9 and the gRNA. Cas9 is a bilobed multidomain protein composed of the REC lobe whose function assures correct binding of the gRNA to the DNA target, and the NUC lobe which contains the two nuclease domains (HNH and RuvC) as well as the PAM interacting domain. These two lobes are connected with a bridge helix (BH). Cas9 is guided by the gRNA, which is composed of the crRNA, complement to the target sequence, and the tracrRNA scaffold connecting with Cas9. Upon recognition of the PAM sequence by the formed heteroduplex gRNA:DNA:Cas9, Cas9 nuclease domains ultimately cleave the DNA target 3 nt upstream the PAM.

The CRISPR revolution started from a seminal work from Jinek and colleagues showing that the crRNA can be programmed to target a specific sequence *in vitro*, and it can be linked with the tracrRNA into a single RNA element (sgRNA, interchangeably used with gRNA) (24) (Fig. I-3). Later in January 2013, two independent labs, namely George Church's and Feng Zhang's labs, showed that these gRNAs could be used for genetic engineering human cell lines (47, 61). When targeted to a living organism's genomic DNA, the "blunt" cut initiates the endogenous repair mechanisms, such as HR or non-homologous end joining (NHEJ) and thus offers a window to perform genome engineering (Fig. I-4, Box 1).

The attractiveness of CRISPR technology also lies in the engineering versatility offered on both the nuclease and the guide, as well as its diverse catalogue of *cas* genes provided by the different CRISPR containing living organisms (62). The structural characterization of Cas9 has led to the development of mutagenized variants with various catalytic properties, e.g. the inactivation of one of the nuclease activity domains ($RuvC^{D10A}$ or HNH^{H840A}, Cas9n) to perform a single-strand DNA break (nick); modification of Cas9's PAM-interacting domain to increase the recognition range PAM sequences, e.g. Hu and colleagues evolved Cas9 to recognize NG, GAA and GAT as PAM sequences (63); or modifications reducing the likelihood of interaction between the gRNA and an off-target genomic sites (64, 65). Complete disruption of the two nuclease activities ($RuvC^{D10A}$ and HNH^{H840A}) results in a catalytically inactive Cas9 or dead-Cas9 (dCas9) which has been exploited as a mean of interfering with transcriptional regulation by sterically hindering the transcriptional activity of the RNA polymerase (66, 67). Additionally, dCas9 can be fused to transcription factors such as the mammalian transcriptional repressor domain Mxi1 to

further improve the efficiency of repression, or with the tripartite VP64-p65-Rta (VPR), composed of three different transcription activators, for gene activation (Figure I-5) (68–72). Finally, other kind of CRISPR nucleases such as Cas12a (also known as Cpf1) from *Franciscella novocida*, and the recently identified CasX (also known as Cas12e) from *Deltaproteobacteria bacterium*, whose small size and different PAM recognition specificities ultimately make them established alternatives to Cas9 (73–75).

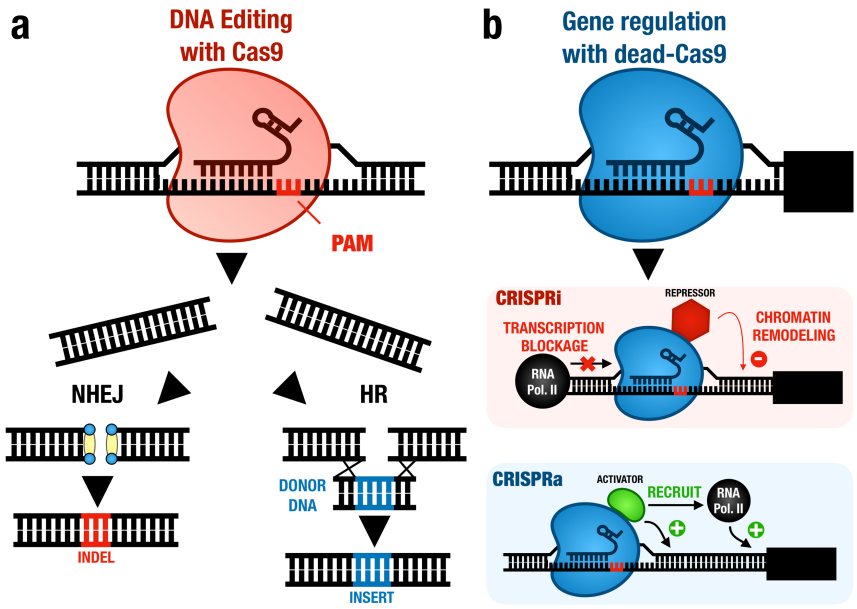


Figure I-4. CRISPR/Cas9 genome editing and gene regulation. a. Genome editing using CRISPR/Cas9 technology**. b.** Gene regulation using dCas9-based technologies.

Thus, the programmability of the gRNA allows the user to target virtually any position in the genome of interest, considering the presence of a PAM sequence. However, correct functionality of the gRNA is also subject to several parameters depending on the desired engineering approach, e.g. improper GC content can affect the stability of the RNA secondary structure (76), or targeting a non-unique region might create unintended binding elsewhere in the genome, called off-targets (77). For most of CRISPR/dCas9-based applications, the distance to the TSS affect has been characterized as an important factor for proper regulation, e.g. -100 to 0bp for downregulation. Nucleosomes, the main component of the chromatin structure, as well as transcription factors can also sterically hinder proper binding of the gRNA (78). Further modifications on the gRNA are usually introduced in the spacer, e.g. to improve on-target specificity, or in the scaffold by adding extra RNA stem-loops, e.g. self-processing elements to prevent degradation. Additionally, processing elements are also exploited for multiplexing by collocating these cleavable elements between each gRNA sequence, which I will further detail in **Chapter IV**.

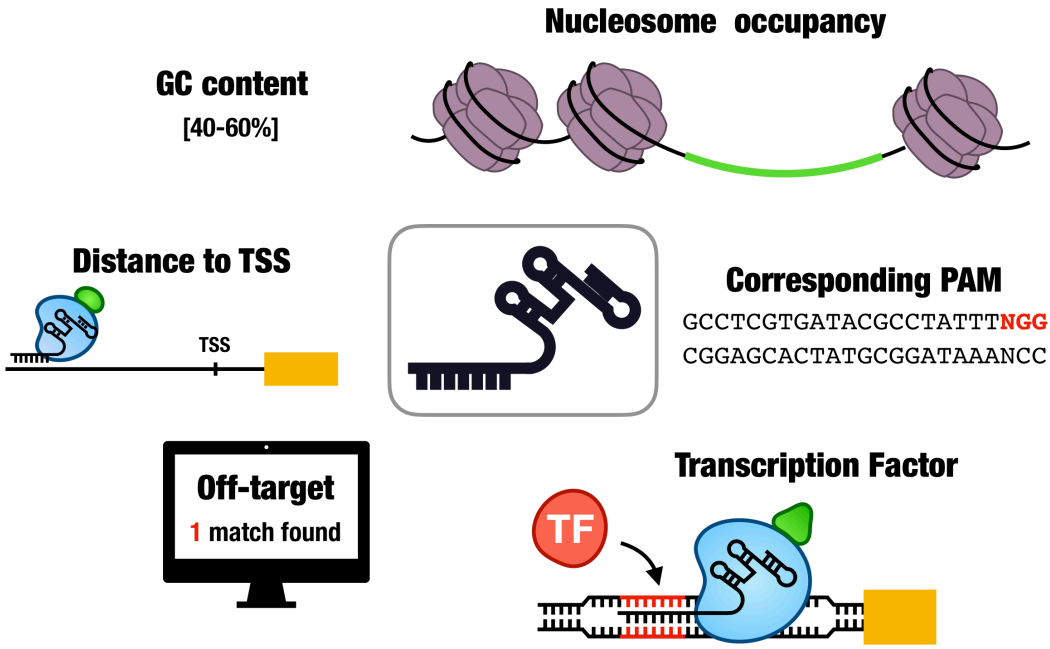


Figure I-5. Key features considered for gRNA design. GC content affects RNA secondary structure which can impede proper binding to the DNA target or to Cas9. Depending on the nuclease employed, the presence of a PAM in the target is critical for genome editing to occur. Off-target analysis which is routinely performed to prevent unwanted binding in the genome. For most of CRISPR/dCas9-based applications, the distance to the TSS affect has been characterized as an important factor for proper regulation, e.g. -100 to 0bp for downregulation. Nucleosome occupancy as well as transcription factors can influence proper binding of the gRNA.

1.3 Engineering *S. cerevisiae* for the production of oleochemicals

Quintessential cell factory

Throughout its extensive utilization, *S. cerevisiae*'s metabolism has been significantly characterized, and it is, consequently, one of the favored microbes for the production of many valuable compounds (79–81). It was the first eukaryotic organism to be sequenced and extensive knowledge regarding its metabolism and physiology is available into several databases (21). Based on this knowledge, the reconstruction of its cellular metabolism has been achieved on the genome scale, generating so-called GEMs, which ultimately allows quantitative predictions of metabolism, e.g. estimating the growth rate under certain culture conditions or simulating gene deletions (39, 80). Additionally, numerous pathways also occurring in higher eukaryotes such human and plant are highly conserved in yeast, which not only has allowed to elucidate the molecular mechanisms of these pathways, e.g. the discovery of autophagy in *S. cerevisiae*, but also facilitated the expression of heterologous genes from these organisms (82, 83). It has a strong track-record for large-scale industrial cultivation, due to its short generation time (doubling time of approximately 2 h at 30°C), and robustness to harsh environments such as low pH and resistance to phages, known to frequently fault bacteria-based cultivations. It naturally possesses an extremely efficient

homologous recombination system (HR) which greatly facilitates genome engineering such as the insertion of multiple genetic elements. Significant scientific contributions also arose from the use of the yeast collections, e.g. a yeast library carrying each individual gene knockout, particularly in screens for synthetic lethality and investigating novel drug targets (84–86).

Anthropocene era – the undesired exponential

My graduate studies have sorrowfully coincided with records of heatwaves that stormed Europe, with France hitting its all-time temperature record in June 2019 with 46°C in the southern village of Gallargues-le-Montueux (87). Concentrations of greenhouse gases (GHGs) — including CO_2, N_2O, and CH_4 — in the atmosphere have drastically increased since the beginning of the industrialization of our society, primarily coming from our dependence on fossil sources (88). Experts' reports regarding the forecast on the climate are unequivocal: inaction on the rising global GHG emissions will be ineluctably followed by increased temperature at higher frequency (89). The global increase of temperature is currently devastating several ecosystems, and recent predictions have placed 1.000.000 species threatened with extinction, with Bramble Cay melomys (*Melomys rubicola*) being the first mammal recognized as extinct due to climate change (90–92). Additionally, the permafrost region which holds twice more carbon than the atmosphere could not only collapse as the temperature rises, but would also transition from a white reflecting surface region to a dark absorbing one, ultimately creating a snowball effect (93, 94). Thus, substantial efforts are needed to prevent reaching the "2°C above pre-industrial era" threshold established at the Conference of the Parties to the United Nations Framework Convention on Climate Change (COP21) in 2015.

Production of oleochemicals in S. cerevisiae

In order to palliate the harmful effect of climate change, there has been an increasing interest in exploiting renewable resources for the synthesis of oleochemicals (95) (Fig. I-4). Oleochemicals are described as a class of hydrocarbon compounds, i.e. composed of carbon and hydrogen, generally extracted from vegetable lipids (96). These lipids are constituted of fatty acids (FAs) esterified with different backbones, with triacylglycerols (TAGs) - three FA units esterified to a glycerol backbone - as the main component (97). Depending on the origin, chemical processing of these lipids into basic oleochemicals can yield different products such as free fatty acids (FFAs), fatty alcohols, alkenes, or alkenes, whose value depends on their acyl chain length, saturation level, terminal groups, and applications (98).

While oleochemicals offer a cost-effective substitute for petrochemicals and fuels, important sustainability and technical challenges from their utilization remain to be addressed. First, many of these oleochemicals derive from the extraction of lipids from exotic plant species that only grow in specific regions on Earth, which ultimately creates

economic instabilities as supply chains are controlled by few suppliers; e.g. 85% of the world palm oil comes from the Malay Archipelago (99). Also, driven by the increasing demand for palm oil, large expansion of the land used to produce it has led to significant deforestation and biodiversity loss, e.g. annual conversion of 270,000 ha of tropical forest conversion to palm trees from 2000–2011 in the main exporting countries (100). Finally, engineering these plants remains extremely challenging due to their slow growth and the lack of efficient tools. Altogether, efforts to alternatively produce oleochemicals by yeast and other microbes through metabolic engineering have increasingly been investigated in the recent years (98, 101, 102) (Fig. I-6).

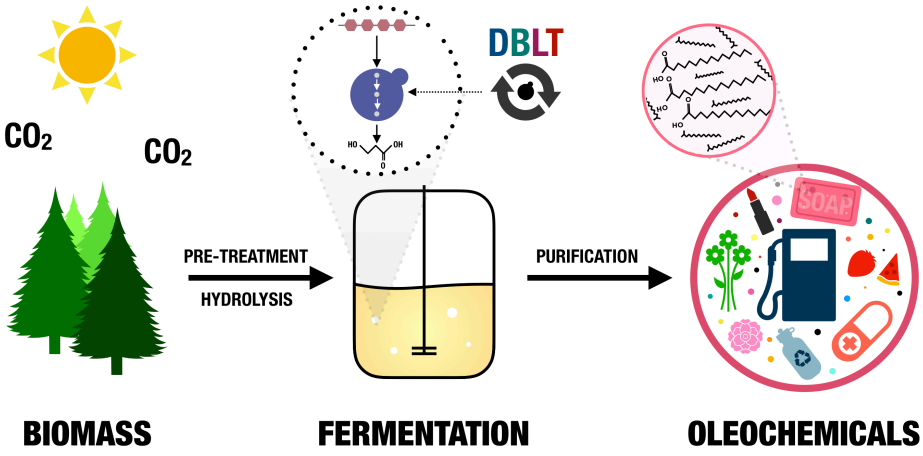


Figure I-6. Biorefinery for the production of oleochemicals. Pre-treatment and hydrolysis of biomass resources such as corn, grass or wood, yield sugars that can be fermented by microbial cell factories for generating high-value oleochemicals. Applications of oleochemicals range transportation fuels, cosmetics, to lubricants and bioplastics.

Since oleochemicals are lipid-derived products, common metabolic engineering strategies applied for microbial cell factories often focus on redirecting carbon flux from sugar to the fatty acid metabolism (Fig. I-7). In *S. cerevisiae*, the synthesis of FAs is catalyzed by the fatty acid synthase (FAS) complex, encoded by the *FAS1* and *FAS2* genes, where malonyl-CoA and acetyl-CoA are used as building blocks to form acyl-CoAs (103). Malonyl-CoA is synthesized from acetyl-CoA by acetyl-CoA carboxylase (Acc1), an enzyme tightly regulated at the transcriptional level by Ino2/4 and Opi1, as well as at the post-transcriptional level by Snf1, which altogether makes this reaction the rate limiting step of the FA biosynthesis (104–106). Malonyl-CoA is iteratively utilized by the FAS complex as an elongation block to extend the size of the acyl-chain. In *S. cerevisiae*, most of the acyl-CoA generated by the FAS complex is naturally composed of long chain fatty acids (LCFAs; C16 and C18), complemented by a small fraction composed of C12 and C14. These FAs can be further elongated to very long chain fatty acids (VLCFAs; C20 up to C26), carried out by elongases (Elo1, Elo2, and Elo3), located at the endoplasmic reticulum

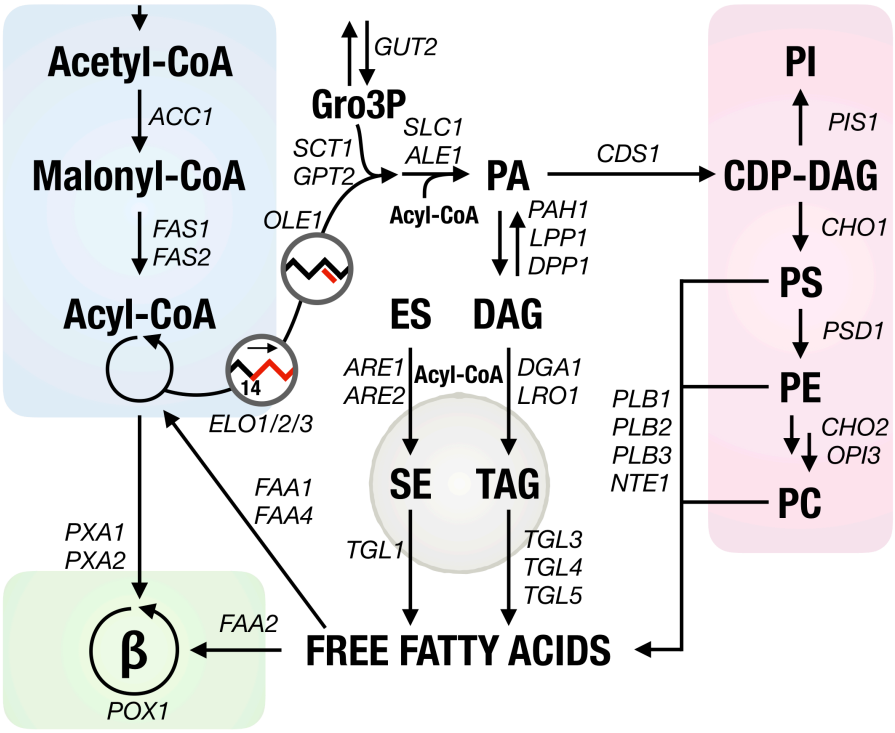


Figure I-7. Fatty acid biogenesis and lipid metabolism in *S. cerevisiae*. Metabolites. Gro3P: glycerol-3-phosphate; PA, phosphatidic acid; CDP-DAG: cytidine diphosphate diacylglycerol; PC: phosphatidylcholine; PE: phosphatidylethanolamine; PI: phosphatidylinositol; PS: phosphatidylserine; ES: ergosterol; SE: sterol esters; DAG: diacylglycerol; TAG: triacylglycerol; β: β-oxidation. **Gene names.** *ACC1:* acetyl-CoA carboxylase; *FAS1/2*: fatty acid synthase; *FAA1/4:* fatty acyl-CoA synthetases, *SCT1* and *GPT2*: glycerol-3-phosphate O-acyltransferase; *GUT2*: glycerol-3-phosphate dehydrogenase; *ALE1*: lysophospholipid acyltransferase; *SLC1*: 1-acyl-sn-glycerol-3-phosphate acyltransferase; *POX1*: fatty acyl-CoA oxidase, *DGA1*: diacylglycerol acyltransferase, *LRO1*: phospholipid:diacylglycerol acyltransferase, *ARE1* and *ARE2*: acyl-CoA:sterol acyltransferases; *PXA1/2*: subunit of peroxisomal ABC-transporter; *ELO1/2/3*: fatty-acid elongases; *TGL3/4/5*: TAG lipases; *TGL1*: sterol esterase; *PLB1/2/3*: phospholipases B; *NTE1*: serine esterase; *PAH1*: phosphatidic acid phosphohydrolase; *CDS1*: phosphatidate cytidylyltransferase; *DPP1*: diacylglycerol pyrophosphate phosphatase; *LPP1*: lipid phosphate phosphatase; *PIS1*: CDP-diacylglycerol-inositol 3-phosphatidyltransferase; *CHO1*: phosphatidylserine decarboxylase proenzyme; *CHO2*: phosphatidylethanolamine N-methyltransferase; *OPI3*: phosphatidyl-N-methylethanolamine N-methyltransferase

(ER), as well as desaturated at the Δ9 position by the only fatty acid desaturase Ole1 (Fig. I-7) (107–110). Since the acyl chain length and saturation level of oleochemicals determine their value and application, several strategies to tweak the abundance of the different FAs have been relying on fine-tuning these genes, e.g. overexpression of *ELO2* to accumulate C22 FAs (111), engineering the endogenous FAS genes to indrocuce heterologous parts which ultimately favored the generation of short-chain FAs (SCFAs) (112), or increasing the rate of C18:1 by replacing *OLE1* by the Δ9 desaturase gene, *ChDes9-2*, from *Calanus hyperboreus* with higher substrate specificity for C18:0 (113).

Intracellular pools of acyl-CoAs are then either used as substrates for generating major lipid classes in the cell, namely phospholipids (PLs), TAGs, sterol esters (SEs), and sphingolipids, or, imported to peroxisomes for degradation to acetyl-CoA through the β-oxidation pathway (Fig. I-7) (114). For PLs and TAGs, acyl-CoAs are sequentially esterified into a glycerol backbone via different acyl transferases, producing phosphatidic acid (PA) (Fig. I-7) (115). PA is a key intermediate of yeast glycerophospholipid biosynthesis as it can either be metabolically converted to diacylglycerol (DAG) for the biogenesis of TAGs, or to cytidine diphosphate diacylglycerol (CPD-DAG), precursor of the different PL classes. SEs synthesis is catalyzed by the acyl-CoA:sterol acyltransferases, encoded by *ARE1* and *ARE2* (116). The hydrolysis of the major lipid classes involves lipases and esterases, and results in a cleavage of the esterified fatty acyl-chain into FFAs (115, 117). These FFAs are directly converted to acyl-CoA by one of the fatty acyl-CoA synthetases, mainly Faa1 and Faa4 in the cytosol, and Faa2 for degradation through β-oxidation in the peroxisome (Fig. I-7).

Besides FFAs and TAGs, synthesis of oleochemicals is not feasible endogenously in *S. cerevisiae*. Thus, efforts have been made, these recent years, towards identifying critical heterologous enzymes catalyzing reactions producing these downstream products (98, 102). Fatty alcohols are used in numerous industrial applications, e.g. surfactants, detergents, and cosmetics (Fig. I-6). Production of fatty alcohols in *S. cerevisiae* can efficiently be achieved from FFAs in a two-step reaction, a carboxylic acid reductase, e.g. from *Mycobacterium marinum*, converts FFAs to fatty aldehydes which can be then converted to fatty alcohols using the endogenous alcohol dehydrogenase (e.g. Adh5) (Fig. I-8). This strategy is usually coupled to removing genes responsible for the degradation of the intermediate metabolites, i.e. the main acyl-CoA activating genes (*FAA1* and *FAA4*) as well the first step of the beta-oxidation (*POX1*) for FFA degradation, and the fatty aldehyde dehydrogenase which converts fatty aldehydes back to FFAs (*HFD1*) (Fig. I-8) (118–120). Another route can be achieved using acyl-CoAs as main substrate through fatty acyl-CoA reductases (FAR) enzymes (121). As mentioned above, fatty acid metabolism is tightly regulated, which ultimately limits metabolic engineering strategies focusing on improving acyl-CoA pools, in comparison with FFAs which can accumulate to high levels (200 fold higher than fatty acyl-CoA) by the removal of *FAA1* and *FAA4* (119).

Fatty alcohols can be esterified with acyl-CoAs by wax ester synthases to produce wax esters (WEs) (Fig. I-8) (121–123). WEs are traditionally extracted from the seed oil of *Simmondsia chinensis* (jojoba) and are applied in a broad range of industrial applications, e.g. cosmetic, anti-inflammatory ointment, and candles (124). For the production of WEs in *S. cerevisiae*, since fatty acids present in the jojoba oil are mostly composed of C20:1 and C22:1, recent studies have extensively focused on obtaining the required acyl chain length by overexpressing *ELO2*, to push from C16/18 to C20/22, and removing *ELO3*, to prevent elongation to C24/26, together with expressing heterologous FAR with higher substrate affinity towards these long chain fatty acids (111, 125, 126). As with fatty alcohols, alkene and alkanes can be derived from fatty aldehydes through fatty aldehyde deformylating oxygenases, or from acyl-CoAs using decarboxylases (127, 128). In contrast

to WEs, alkanes and alkenes usually require shorter acyl chain (C7 to C13) for industrial application, which is challenging to accumulate in *S. cerevisiae* due to the fatty acid machinery predominantly generating C16 and C18 fatty acid precursors (129). Recent studies have enabled the synthesis of medium chain alkanes and alkenes in *S. cerevisiae* by primarily engineering endogenous FAS genes into a single enzyme with domains borrowed from other fungal species, yielding a FAS complex generating short and medium chain fatty acids (112, 130).

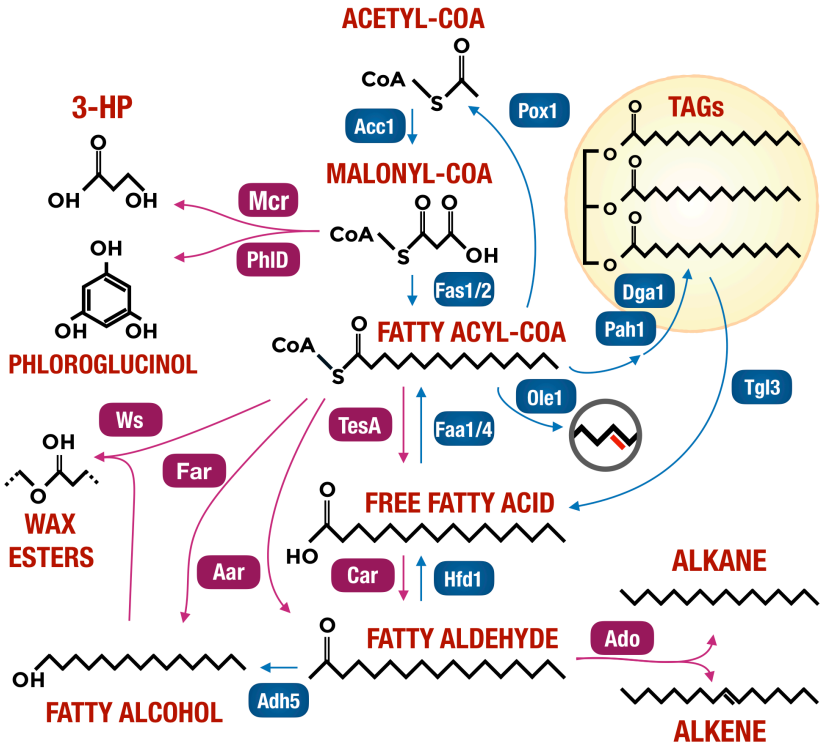


Figure I-8. Strategies for the production of oleochemicals in *S. cerevisiae*. Acc1: acetyl-CoA carboxylase 1; Fas1/2 : fatty acid synthase 1 and 2; Pox1: fatty-acyl coenzyme A oxidase; Mcr: malonyl-CoA reductase; PhlD: type III polyketide synthase; Ole1: fatty acid monodesaturase; TesA: acyl-CoA thioesterase I; Faa1/4: fatty-acid activation 1 and 4; Pah1: phosphatidic acid phosphohydrolase 1; Dga1: diacylglycerol acyltransferase 1; Tgl3: triacylglycerol lipase 3; Ws: wax ester synthase; Far: fatty acyl-CoA reductase; Aar: acyl-ACP reductase; Hfd1: Fatty aldehyde dehydrogenase; Ado: aldehyde-deformylating oxygenase; Adh5: alcohol dehydrogenase 5. Blue arrows: endogenous enzymes from *S. cerevisiae*, in purple: heterologous enzymes.

Other engineering strategies include the engineering of permeases to secrete toxic oleochemicals, e.g. fatty alcohols (131–135). Also, since *S. cerevisiae* is an eukaryote, subcellular compartmentalization of complete metabolic pathways can be achieved in different organelles such as the mitochondria, ER, and peroxisomes, with increased efficiency due to the availability of certain metabolites, e.g. unique presence of heme in the mitochondria (136). Besides acetyl-CoA and malonyl-CoA, FA synthesis consumes large amounts of NADPH, whose main cellular formation takes place in the oxidative pentose

phosphate pathway (PPP) (107). Thus, examples of recent strategies have rewired fluxes from glycolysis through the PPP to generate more NADPH, as well as expressing enzymes regenerating NADPH, e.g. $NADP^+$-dependent glyceraldehyde-3-phosphate dehydrogenase (GAPN) from *Bacillus cereus* which converts 3-P-glyceraldehyde to 3-P glycerate in a single reaction using $NADP^+$ as co-factor (137, 138).

1.4 Scope of the thesis

In **Chapter I**, I have introduced the concept of metabolic engineering metabolism and how it can be purposely modified to enhance the production of metabolites of interest. The DBTL cycle which leverages tools from systems biology and synthetic biology, is commonly implemented to develop efficient cell factories. Building a strain includes introducing genetic modifications such as gene deletion, overexpression, and exact regulation to improve pathway efficiency and product yield. In this context, CRISPR has in the recent years been adapted for manipulating a broad range of living organisms, including their metabolism. This chapter has also been focusing on giving an overview of CRISPR technology from its discovery to its application in living organisms (**Paper I**). This thesis will explore how CRISPR technologies can be applied in the context of building efficient yeast cell factories for the production of oleochemicals.

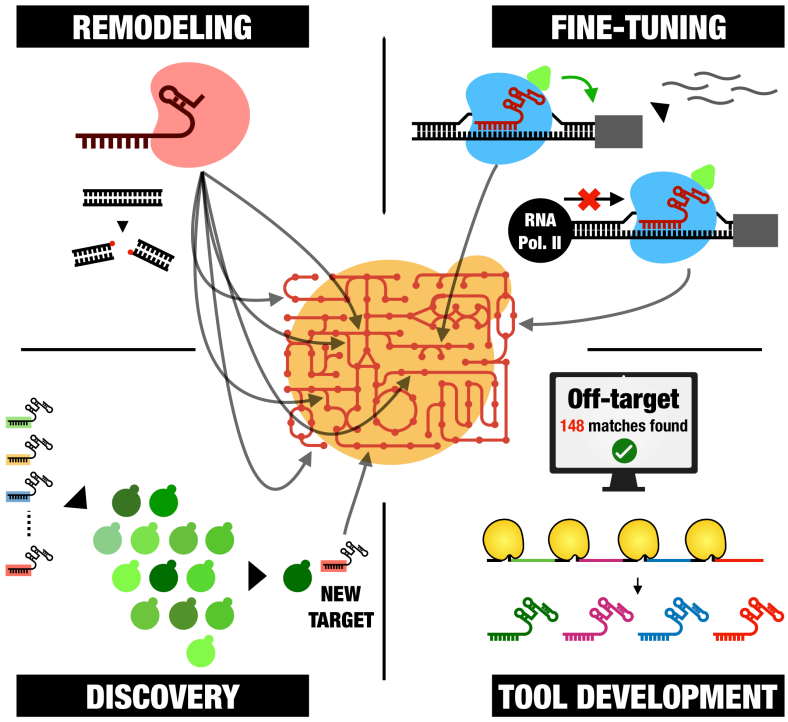
In **Chapter II**, I will present two studies underlining the time-efficiency of CRISPR technology for remodeling *S. cerevisiae*'s lipid metabolism for increasing the production of free fatty acids and triacylglycerols, respectively. I will present results from **Paper II**, in which we generated a free fatty acid production strain by minimizing the lipid metabolism network and deregulating different feedback loops. Then results from **Paper III** will be presented, where several metabolic fluxes were engineered to overcome the main limiting steps in fatty acid biosynthesis, ultimately generating a strain accumulating the highest levels of triacylglycerols reported in *S. cerevisiae* so far.

After focusing on genome remodeling by principally removing multiple genes involved in a particular pathway, **Chapter III** will explore dCas9-based technologies for fine-tuning of biosynthetic pathways, i.e. precise control over gene expression. This is often a key step for the efficient synthesis of a particular target compound but is relatively challenging to achieve through traditional engineering strategies. The CRISPR system, besides its efficient endonuclease activity, can enable transcriptional regulation *via* the catalytically deactivated form of the Cas9 protein coupled to transcription factors. This chapter will be supported by two studies: In **Paper IV**, we benchmarked two dCas9-mediated systems for controlling the expression of 14 genes targeted by a total of 101 gRNAs. We also demonstrated the impact of single and multiplex gRNA strategies for reprogramming expression of multiple genes in the isoprenoid and TAG biosynthetic pathways. In **Paper V**, we proposed a framework based on computational techniques for predicting key genes that potentially favor the production of the endogenous metabolites, and on the versatility of CRISPR technology for fine-tuning gene expression coupled with an intracellular

biosensor for selecting high-producing cells. This was carried out by designing targeted gRNA libraries to explore multiple metabolic pathways.

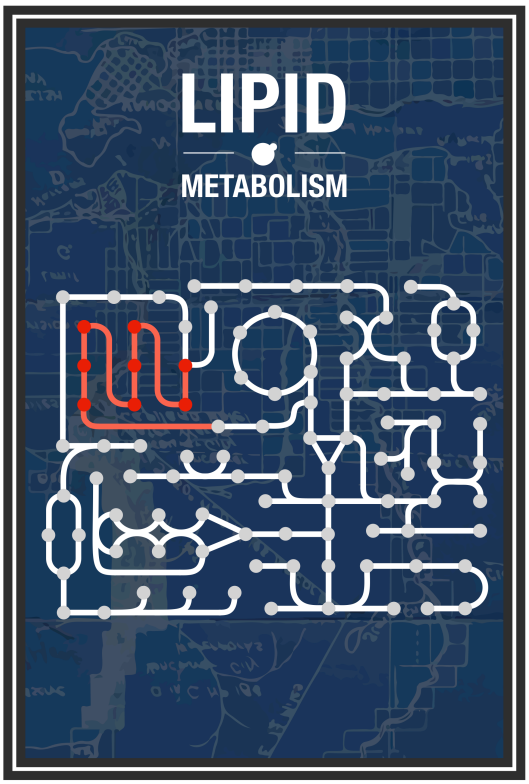
In **Chapter IV**, we seek to expand the CRISPR repertoire by building new tools to accelerate microbial cell factory design. This chapter is supported by several studies: *(i)* exploiting CRISPR-associated endoribonuclease Csy4 for multiplex genome engineering and transcriptional regulation *via* processing an RNA transcript into multiple gRNAs (**Paper VI**), *(ii)* using promiscuous on-target score of a given gRNA which ultimately allows to target multiple loci at once (**Paper VII**), and *(iii)* characterization of other CRISPR nucleases.

Finally, **Chapter V** will be the concluding chapter of the thesis focusing on the future perspectives of CRISPR technology in engineering metabolism and its applications in yeast. This section will emphasize the recent technologies that have been developed and are yet to be applied to yeast metabolism. Additionally, comments will be made regarding the future utilization of *S. cerevisiae* as chassis for metabolic engineering and advancing CRISPR technologies.



Graphical abstract of the thesis.

Chapter II: Remodeling specific metabolic pathways



2.1 Towards a minimal lipid metabolism (Paper II)

As previously mentioned, FFAs are attractive endogenous chemicals that can readily be converted into a broad range of industrial compounds (Fig. I-5). These metabolites are mostly extracted from plant oils after hydrolysis of TAGs, and engineering efforts towards a more sustainable process through the development of microbial cells capable of efficiently converting sugars into FFAs are actively being sought. Fatty acid synthesis is strongly regulated by feedback mechanisms and FFAs are found natively only in low levels (less than 0.1% of dry cell weight) (114). These challenges make, theoretically, *S. cerevisiae* an unfavored chassis for the production of these compounds. In **Paper II**, harnessing CRISPR efficiency, we sought to investigate the dynamics of lipid metabolism surrounding FFA biosynthesis by minimizing *S. cerevisiae*'s lipid metabolic network towards its essential components, i.e. fatty acid and membrane lipid biosynthesis.

In *S. cerevisiae*, the main routes to generate FFAs come from hydrolysis of the main lipid classes, i.e. TAG lipases, e.g. *TGL3*, *TGL4*, and *TGL5*, degrading TAGs into DAGs and FFAs; sterol esterase, e.g. *TGL1* and *YEH1*, degrading SEs into sterols and FFAs; and phospholipases, e.g. *PLB1*, *PLB2*, and *PLB3*, which cleave the acyl-chain from phospholipids (139–143). As mentioned above, FFAs derived from hydrolysis of the main

lipid classes, are directly converted to acyl-CoA by fatty acyl-CoA synthetases which keeps FFA levels in the cytoplasm low. Previous engineering strategies have shown that deletion of the two main acyl-CoA synthetase genes *FAA1* and *FAA4* results in de-regulation of fatty acid biosynthesis and over-accumulation of FFAs orders of magnitude above a wild-type strain. This double deletion is usually the starting platform strain for metabolic engineering of *S. cerevisiae* for the production of FFAs. To further prevent FFA and acyl-CoA degradation and provide a more constrained lipid metabolism network, deletion the fatty acyl-CoA oxidase encoded by *POX1*, which catalyzes the first step of the β-oxidation pathway located in the peroxisomes results in a strain (RP02) accumulating to 53 mg·gDCW^{-1} of FFAs, significantly higher than wild-type levels.

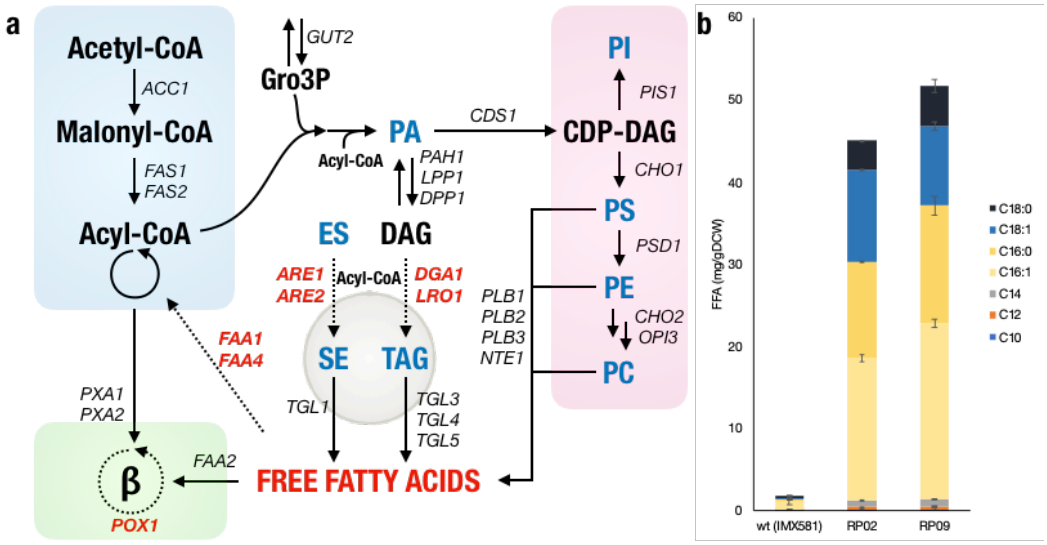


Figure II-1. Simplified lipid metabolism fluxes and genes in *S. cerevisiae* and the effect of disrupting lipid droplet formation on FFA biogenesis. a. The lipid species experimentally quantified in blue: PA, phosphatidic acid; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol; PS, phosphatidylserine; SE, sterol esters; TAG, triacylglycerols; with Free Fatty Acids in red as it is the main metabolite studied. Genes highlighted in red are targeted for deletion in the RP09 strain: fatty acyl-CoA synthetases *FAA1* and *FAA4*, fatty acyl-CoA oxidase *POX1*, diacylglycerol acyltransferase *DGA1*, phospholipid:diacylglycerol acyltransferase *LRO1*, and acyl-CoA:sterol acyltransferases *ARE1* and *ARE2*. **b.** Total FFAs were quantified for the wild-type wt, RP02 strain lacking fatty acid reactivation (*FAA1* and *FAA4* deletions), β-oxidation (*POX1* deletion), and RP09 with additionally deleted storage lipid formation (*DGA1*, *LRO1*, *ARE1*, and *ARE2* deletion).

Next, we investigated the role of the storage-lipid formation pathways in FFA biogenesis. Thus, we deleted the main storage lipid biogenesis genes, *DGA1* encoding diacylglycerol acyltransferase, *LRO1* encoding phospholipid:diacylglycerol acyltransferase, and *ARE1* and *ARE2* encoding acyl-CoA:sterol acyltransferases (Fig. II-1). The resulting strain (RP09: *Δfaa1 Δfaa4 Δpox1 Δdga1 Δlro1 Δare1 Δare2*), did not significantly accumulate more FFAs compared to RP02 (53 mg·gDCW^{-1} and 57 mg·gDCW^{-1} respectively, Student's

t-test p-value > 0.05). Thus, the combinatorial deletion of the storage-lipid formation genes shows that FFAs biogenesis is independent of the lipid storage pathway.

The strain RP09 allowed for further study of the fatty acid dynamics and pathways involved in FFA formation, since it was devoid of neutral lipids and therefore these could be excluded from the fatty acid pools existent in the cell. Here, we hypothesized that removal of storage lipid formation could have led to an accumulation of intermediates such as diacylglycerols (DAGs), which are formed by dephosphorylation of PA. Notably, previous reports have demonstrated that high-levels of PA reduce translocation of the transcriptional regulator Opi1 to the nucleus, which ultimately prevent its binding to Ino2. Since Ino2 is a transcription factor activating several lipid biosynthesis genes, we sought to increase PA levels by removing the dephosphorylation step from PA to DAG. Deletion of the main phosphatidate phosphatases encoded by *PAH1*, *LPP1* and *DPP1*, from RP09 resulting in the strain MLM1.0 (MLM1.0: *Δfaa1 Δfaa4 Δpox1 Δdga1 Δlro1 Δare1 Δpah1 Δlpp1 Δdpp1*), accumulated a total of 102 mg·gDCW^{-1} FFAs, a 98% increase from RP09. Removal of the phosphatidate phosphatases resulted in exclusive use of PA for phospholipid biosynthesis, combined with the deregulation through PA signaling, phospholipid levels were subsequently increased by 8-fold compared to RP09, reaching 30 mg·gDCW^{-1} (Fig 3c).

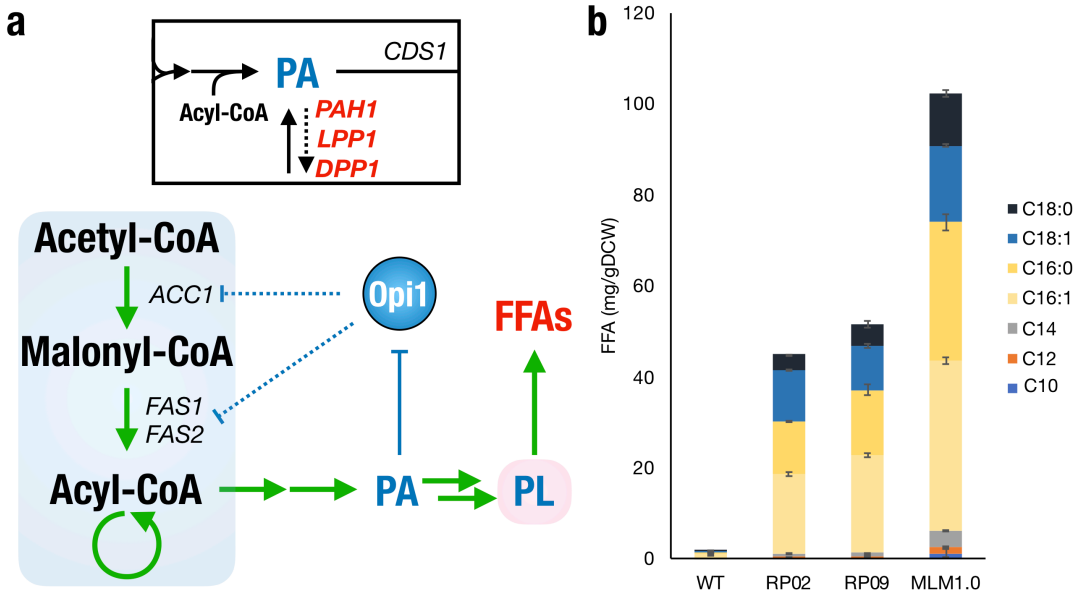


Figure II-2. Effect of deleting the main phosphatidate phosphatases on the FFA levels. a. Deletion of the main phosphatidate phosphatases activates the signaling mechanism linking PA accumulation and the up-regulation of fatty acid biosynthesis through Opi1. PA binds to Opi1, which indirectly causes an up-regulation of *ACC1*, *FAS1*, and *FAS2* genes. **b.** Total FFA quantification of MLM1.0 compared with strains RP09, RP02, and the control strain (wt).

While phosphatidylserine (PS) and cardiolipin (CL) remained unchanged and present in low percentages compared to RP09, a significant accumulation of phosphatidylcholine (PC), phosphatidylethanolamine (PE) and phosphatidylinositol (PI) was observed in MLM1.0 (Figure 3). Notably, we observed that the accumulation of phospholipid was associated with a formation of large membrane structures in MLM1.0 compared to WT, RP02, and RP09 (Fig. II-3), in accordance with previous reports which characterized the expansion of internal ER and nuclear membrane structures upon Opi1 deregulation.

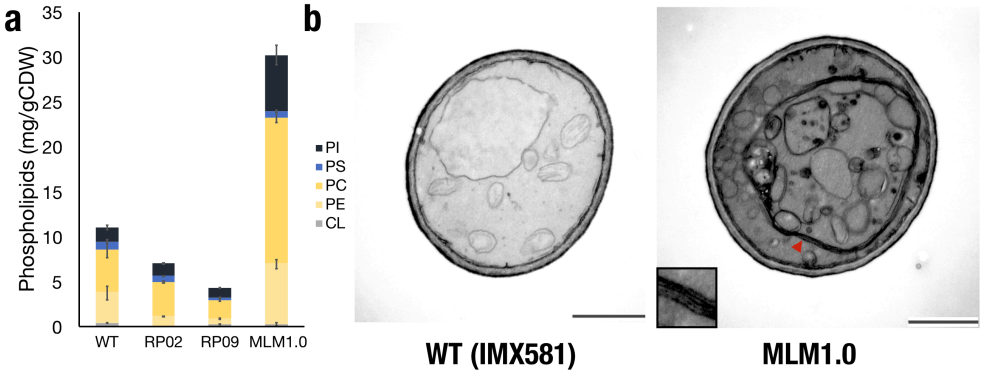


Figure II-3. Phospholipid levels in strains RP09, MLM1.0 and the resulting effect on ER membrane. a. PI, phosphatidylinositol; PS, phosphatidylserine; PC, phosphatidylcholine; PE, phosphatidylethanolamine; CL, cardiolipin. **b.** Transmission electron microscopy of strains generated in this study Observed stacked membrane structures in strain MLM1.0 marked with a red arrow.

Since we observed a combined increase of both FFAs and phospholipids in MLM1.0, we sought to investigate the link between these two lipid classes by targeting the main phospholipase B genes, namely *PLB1* and *PLB2*. These genes encode for enzymes catalyzing the hydrolysis of acyl chains from phospholipids at positions sn-1 and sn-2. Here, deletion of *PLB1* and *PLB2* (MLM1.0 + *Δplb1* + *Δplb2*: *Δfaa1 Δfaa4 Δpox1 Δdga1 Δlro1 Δare1 Δpah1 Δlpp1 Δdpp1 Δplb1 Δplb2*) led to significant decrease of FFAs and an increase of phospholipids by 46% and 105% respectively (p value < 0.05; Student's t-test: one-tailed, two-sample equal variance) (Fig. II-4). This result indicates that phospholipases play an important role in cytosolic and extracellular FFA biogenesis, particularly when the acyl-CoA flux is redirected towards phospholipids.

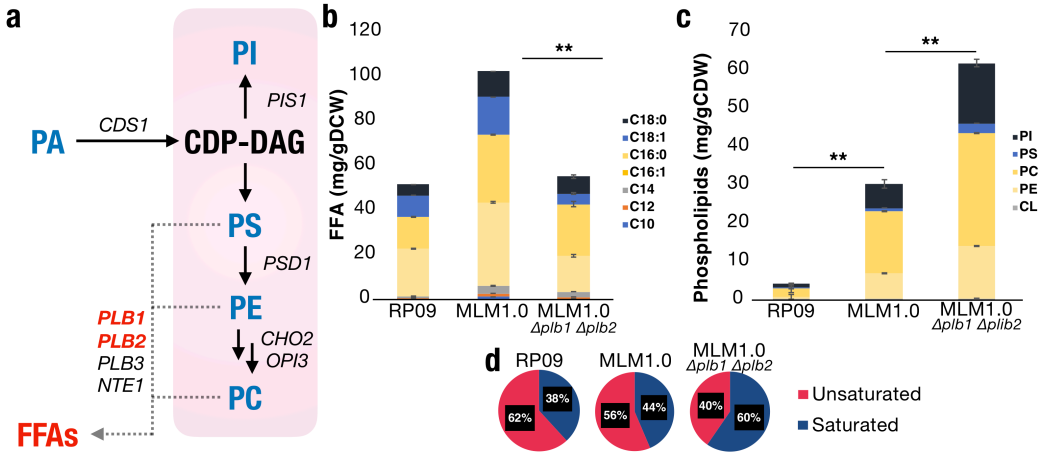


Figure II-4. Effect of deletion of the main phospholipase genes PLB1 and PLB2 on FFA and phospholipid levels. a. Simplified overview of phospholipid synthesis and degradation pathway. **b.** FFA quantification **c.** Phospholipid quantification. **d.** Distribution of FFA acyl classes distribution.

This hypothesis is reinforced by the acyl-chain class distribution composing of the secreted FFAs which showed a decrease of unsaturated FAs (C16:1 and C18:1) from 62% in RP09 to 56% in MLM1.0, as PLs are mostly composed of unsaturated FAs (129).

2.2 "Maximizing" lipid metabolism (Paper III)

As mentioned in **Chapter I**, TAGs are valuable versatile compounds that can be harnessed for biofuel production as well as for nutrition and health (144). TAGs are comprised of three fatty acids esterified with a glycerol backbone, and are stored, together with SEs, in lipid bodies (also referred as lipid droplets (LD)) (115, 117). LDs are composed of a monolayer of phospholipids loaded by several proteins whose roles impact LD biogenesis, cellular distribution, and degradation (145). TAGs serve as an important storage form of energy which, upon requirement, are hydrolyzed by lipases to release DAGs and fatty acids required for membrane synthesis (146). In laboratory condition, *S. cerevisiae* does not naturally accumulate high amounts of TAGs, with levels reaching ~1% of its cell dry weight (145). This is significantly lower than oleaginous species such as *Yarrowia lypolitica* that can accumulate TAGs up to >70% of their dry weight (147, 148).

In this study, in order to increase levels of TAGs we sought to perform three distinct engineering strategies: *(i)* increase fatty acid biosynthesis, *(ii)* push fluxes from PA to TAGs instead of phospholipids, and *(iii)* prevent any degradation of our desired product. For the first two steps, we co-overexpressed genes encoding enzymes catalyzing the first and last steps of TAG biogenesis, namely the conversion of acetyl-CoA to malonyl-CoA by *ACC1* and the conversion of PA to DAG (*PAH1*) and DAG to TAGs (*DGA1*). In that regard, we used a deregulated mutant version of *ACC1*, $ACC1^{S659A/S1157A}$ (*ACC1***), expressed under control of the promoter *HXT7p* that leads to high levels of expression

under low glucose concentrations. We genomically integrated one additional copy of *PAH1* and *DGA1* each, expressed under control of strong constitutive promoters. The resulting strain, ADP (ADP: *HXT7p-ACC1** + PGK1p-PAH1 + TEF1p-DGA1*) accumulated 129 mg·gCDW^{-1} of TAGs, i.e. a 10-fold increase in comparison to the reference strain (Fig. II-2). Expectedly, the accumulation of TAGs also led to an increased number of LDs, as their formation is driven by the synthesis of TAG and SE (Fig. II-6) (117).

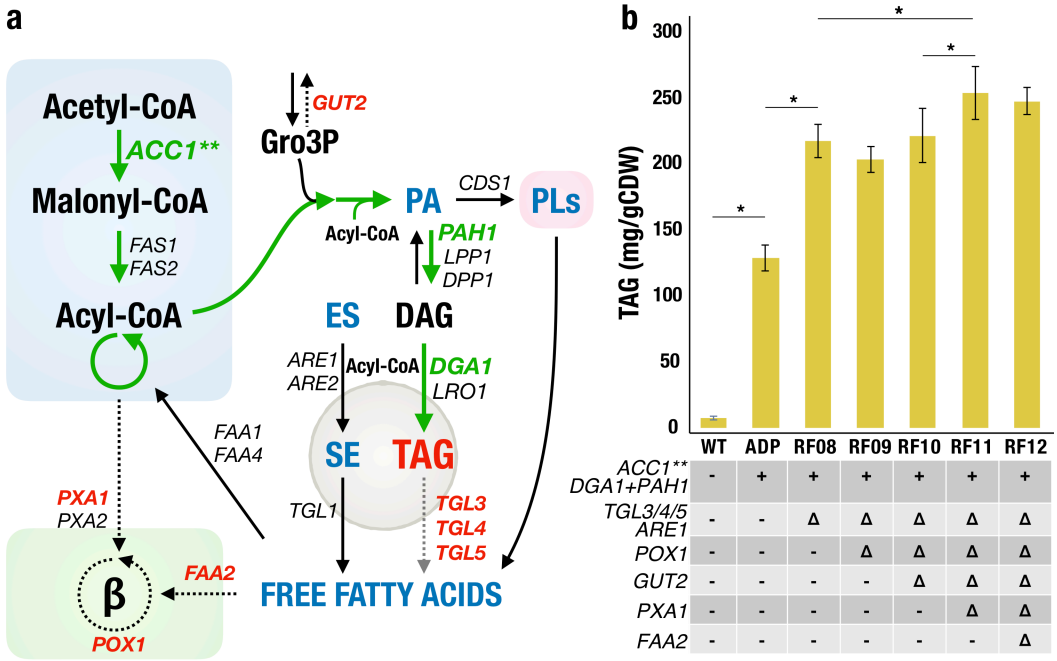


Figure II-5. Metabolic engineering strategy for overproducing TAGs. a. Engineering strategy used for increasing TAG levels. Highlighted in green are genes overexpressed, *ACC1*** (acetyl-CoA carboxylase), *DGA1* (diacylglycerol acyltransferase) and *PAH1* (phosphatidic acid phosphohydrolase). Highlighted in red are deleted genes, *GUT2* (glycerol-3-phosphate dehydrogenase), *ARE1* (acyl-CoA:sterol acyltransferase), *TGL3/4/5* (triacylglycerol lipases), *PXA1* (subunit of peroxisomal ABC-transporter), *POX1* (fatty acyl-CoA oxidase), and *FAA2* (fatty acyl-CoA synthetase). PLs: phospholipids **b.** TAG levels of the engineered strains.

Next, we sought to prevent TAGs from degradation by removing the three main TAG lipases, namely Tgl3, Tgl4, and Tgl5, which has been reported to increase TAG levels (139). Additionally, as neutral lipids stored in LDs are normally composed of 50% TAGs and 50% SEs, we removed sterol ester formation through deletion of *ARE1* in order to obtain a larger fraction of the neutral lipids in form of triacylglycerols. The resulting strain, RF08 (RF08: *HXT7p-ACC1** + PGK1p-PAH1 + TEF1p-DGA1 Δtgl3 Δtgl4 Δtgl5 Δare1*), accumulated 218 mg·gCDW^{-1} TAGs, a 68% increase in TAGs levels compared to the previous strain, ADP (p value < 0.05; Student's t-test: one-tailed, two-sample equal variance) (Fig. II-5). These results indicate a potential direct turnover of TAGs triggered upon their accumulation, in agreement with previous report showing that lipase activity is affected by neutral lipid levels (140).

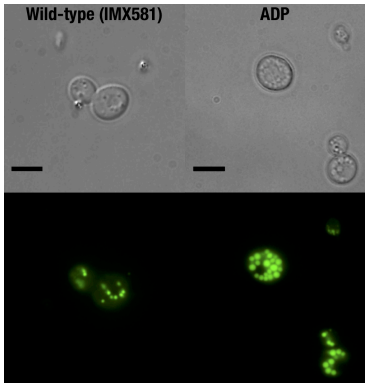


Figure II-6. Lipid droplet staining of strains IMX581 and ADP. Lipid droplets were stained with BODIPY and analyzed with a fluorescence microscope. Black bars represent a 5 μm scale.

Next, we aimed to increase the main constituent of TAGs, namely acyl-CoA and glycerol supplies. For acyl-CoAs, we targeted peroxisomal genes to prevent both FFA and acyl-CoA degradation by deleting *FAA2* and *PXA1*, encoding peroxisomal importers of FFAs and acyl-CoAs respectively, as well as *POX1*. For glycerol backbone supply, we removed the conversion of glycerol-3-phosphate to dihydroxyacetone phosphate by deleting *GUT2*, which encodes the enzyme responsible for this catalytic reaction. Notably, *POX1* removal, which yielded RF09, led to a decrease of 9% in TAGs compared to RF08. Removal of *GUT2*, yielding RF10 strain, accumulated TAGs at a level of 222 mg·gCDW^{-1}, a slight increase compared to RF09, reaching similar levels as RF08 (Fig 2). However, *PXA1* removal, resulting in RF11, significantly improved TAG accumulation, reaching 254 mg·gCDW^{-1}, i.e. a 14% increase in comparison to RF10. Finally, deletion of *FAA2*, resulting in RF12 had no impact on TAG levels (Fig II-5).

2.3 Conclusions of Chapter II

In this **Chapter**, I aimed to highlight Cas9 efficacy for genome remodeling in the context of engineering yeast cell factories for the production of lipid-derived products. Notably, through the engineering strategies applied in these two papers, we learnt that the complexity of the lipid metabolism, i.e. its numerous feedback regulation and balanced homeostasis between the different lipid classes, can be exploited to accumulate some of its compounds. In **Paper II**, we investigated the pathways involved in FFA biogenesis and showed the dynamics through which these metabolites can be biosynthesized. We constrained the acyl-CoA fluxes to phospholipids in absence of fatty acid activating genes (*FAA1* and *FAA4*) and ultimately generated a relatively highly engineered strain, accumulating high levels of FFAs and phospholipids. We exploited the Ino2/Ino4/Opi1 system which regulates the fatty acid and phospholipid synthesis pathways to further increase FFAs levels (149). These results are in agreement with previous reports where *Δpah1* was shown to modulate phospholipids, TAGs, and FFA levels (150). While we mainly focused on FFA production in this study, the titers reported remained below previous reports (119, 151, 152). Notably,

MLM1.0 accumulates high-levels of phospholipids which could also be used as excipients for pharmaceutical applications (153). Additionally, we attempted to take advantage of this pool of phospholipids to further increase FFAs by overexpressing phospholipase B genes, namely *PLB1*, *PLB2* and *PLB3*, but FFAs were not affected by these upregulations (**Paper II**: Figure 5) (154).

For the study on enhancing TAG accumulation, the implemented push-and-pull strategy by overexpressing genes encoding a constitutively active version of *ACC1*, together the last two steps of TAG formation, namely *PAH1* and *DGA1*, ultimately resulted in an accumulation of 12% of the dry cell biomass, more than 10-fold above the levels in the reference strain. Next, the disruption of several lipid droplet associated-proteins involved in both TAG hydrolysis, namely *TGL3*, *TGL4*, *TGL5*, further increased TAG levels up to 22% of the dry cell weight (Fig. II-5). However, further disruption of the beta-oxidation and FFA import in the peroxisome, as well as providing more glycerol-3-phosphate substrate, did not affect TAGs levels. As mentioned above, FFAs are found in low quantity in normal conditions, as they are directly converted to acyl-CoAs by fatty acyl-CoA (114). While not experimentally measured, the disruption of TAG lipases might have further decreased the amount of FFAs available in the cell. Also, theoretically, *FAA2* disruption would have allowed more FFAs to be converted to acyl-CoAs. Previous studies have reported an increase of TAG levels upon *POX1* disruption, which altogether indicates that these disruption strategies, namely the deletion of *FAA2* and *POX1*, could have significantly increase TAG levels only in the presence of TAG lipases (155).

Finally, disruption of the peroxisomal fatty acyl-CoA transporter *PXA1* increased TAG levels up to 25% of the cell dry weight. Since *PXA1* encodes a transporter of fatty acyl-CoA to the peroxisome, I speculate that deleting it ultimately increased the levels of available substrate for TAG biosynthesis, which ultimately points towards a potential limitation in cytosolic acyl-CoA availability for TAG biosynthesis.

Chapter III: Rewiring metabolic fluxes via transcriptional regulation

3.1 Benchmarking CRISPR/dCas9 tools (Paper IV)

As described in **Chapter I**, dCas9 variants can be used for tunable and orthogonal control of gene expression by sterically hindering the transcriptional activity of the RNA polymerase, and by fusing them with repressor and activator domains, such as the mammalian transcriptional repressor domain Mxi1 or with VPR. In this context, these tools can be harnessed for tunable expression of genes encoding enzymes of metabolic pathways to increase productivity. However, at the conception of the study, the metabolic engineering field was lacking robust orthogonal and tunable transcriptional control mechanisms, and metabolic engineers primarily utilized a few characterized promoters for enhancing gene expression. However, CRISPR/dCas9-based systems have, hitherto, not only been reported as robustly achieving several folds of up- or down-regulation. Graded transcriptional patterns can also be achieved depending on where the dCas9 complex is guided in the promoter region.

In **Paper IV**, we benchmarked two dCas9-mediated systems for controlling gene expression of genes targeted by gRNAs in the context of metabolic engineering, i.e. coupled to the production of industrially relevant compounds, namely TAGs and β-carotene (Figure III-1). This project was a collaboration between our lab (SysBio) and Jay Keasling's lab at Danmarks Tekniske Universitet (DTU), where we explored the design, engineering and application of CRISPR-dCas9 technology for multiplex genome

reprogramming of metabolic genes in yeast. The study was one of the first of its kind to describe the systematic use of gRNA libraries and dCas9 designs to explore CRISPR-mediated gene expression tuning for optimizing metabolic fluxes. The first system offers an anhydrotetracycline (aTc) inducible expression of gRNA with dCas9 fused to regulatory domains. The second system combines the gRNA with RNA-stem loops facilitating the recruitment of scaffold-binding domains MCP and PCP fused to either VPR or Mxi, which ultimately allows simultaneous bi-directional and targeted gene regulation (Figure III-1) (156). I was involved, almost exclusively, in characterizing dCas9-VPR/Mxi applied to the lipid metabolism, while my colleagues at DTU were focusing on the scaffolding systems applied to the β-carotene pathway (70).

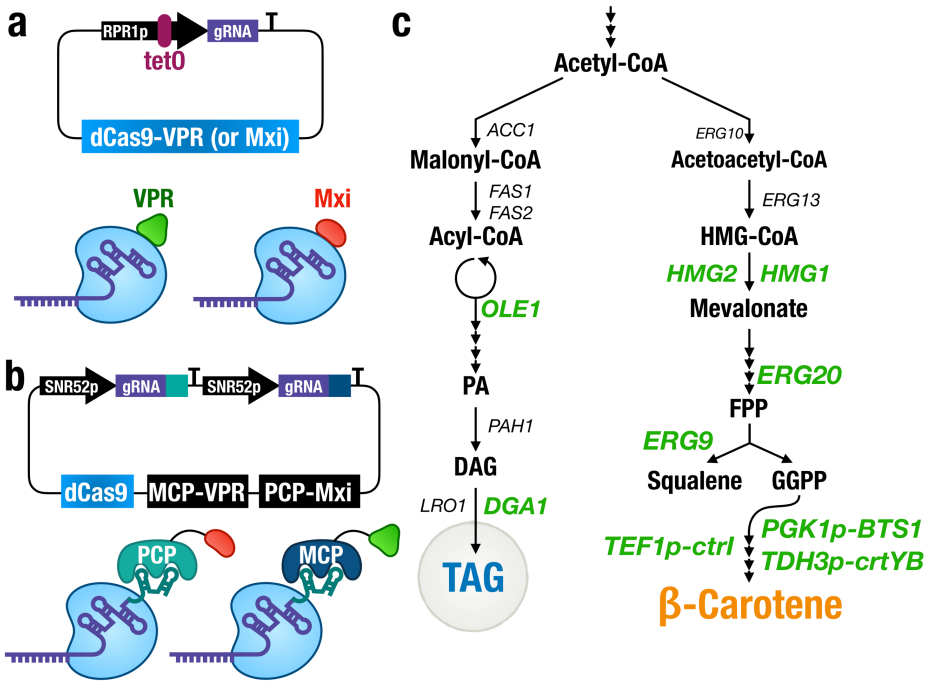


Figure III-1. CRISPR-dCas9 tools benchmarked in β-carotene and TAG biosynthesis pathways. a. aTc-inducible gRNA expression vector with dCas9 fused to VPR or Mxi. **b.** Constitutive gRNA expression vector. The gRNAs are designed with an extra RNA stem-loop to recognize either MCP-VPR or PCP-Mxi ultimately allowing simultaneous repression and activation. **c.** β -carotene and TAG pathways. Marked in green are all the genes targeted in Paper IV for dCas9-based transcriptional regulation

While throughout the study we evaluated more than 100 gRNAs targeting 14 different promoters, the benchmark experiment comparing the two systems was done by targeting two promoters involved in either fatty acid synthesis (*OLE1*) or the mevalonate pathway (*HMG1*) with only a single gRNA per promoter (Fig. III-1). These gRNAs were selected among a small selection of gRNAs targeting the respective promoters retrieved using two webtools (http://chopchop.cbu.uib.no, http://lp2.github.io/yeast-crispri) (Fig. III-2) (157, 158). These tools provided a list of gRNA target sites including important information such

as the distance to TSS or the nucleosome occupancy (Fig. I-5). To assess the efficiency of these gRNAs, we coupled their respective targeted promoter to GFP and monitored the fluorescence level upon gRNA expression. Among the different gRNAs tested, gRNA *OLE1* TSS -381 and gRNA *HMG1* TSS -152 gave the most significant transcriptional regulation when coupled with the inducible and constitutive systems, respectively, and were subsequently selected for benchmarking these dCas9 tools.

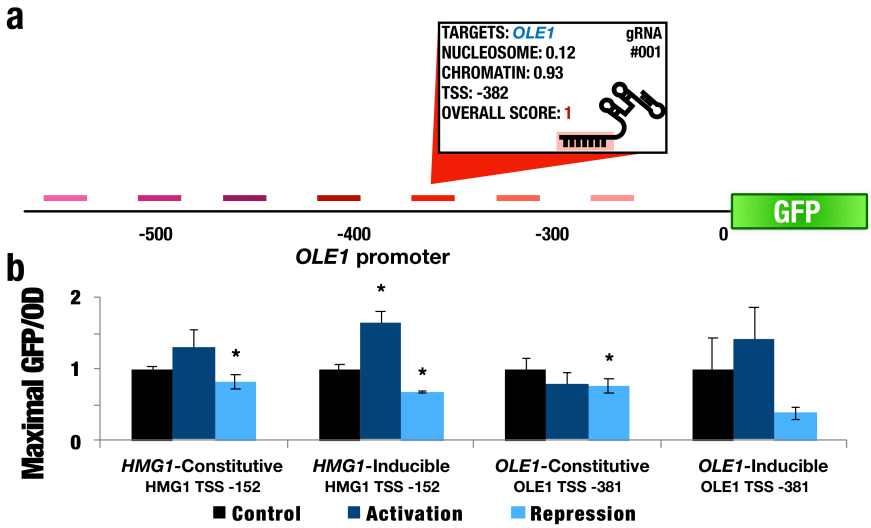


Figure III-2. Transcriptional regulation of gRNAs targeting specific promoters. a. Design of a small gRNA library targeting *OLE1* promoter. The gRNAs were designed to cover different position of the targeted promoter, and characteristics of the position, such as the distance to TSS and the nucleosome occupancy, were also obtained. To assess their transcriptional regulatory effect, the targeted promoters were coupled to GFP. **b.** Fluorescence levels monitored using Biolector. Shown are the maximal fluorescence levels obtained for triplicates grown in synthetic media over the course of 72h.

For the *HGM1* promoter, both the constitutive and inducible system displayed around 50% upregulation and downregulation when coupled with their respective regulator, i.e. downregulation for Mxi and upregulation with VPR. For the *OLE1* promoter, the inducible system inferred around 50% activation and repression using dCas9-VPR and dCas9-Mxi (Fig. III-3, Fig. III-4). Based on these only two promoters, we concluded that the inducible system performed slightly better than the constitutive system. To further test the fine-tuning applicability of the inducible system, we characterized more gRNAs targeting the *OLE1* and *DGA1* promoters, two genes shown to increase TAG biosynthesis when overexpressed (113, 121, 159, 160).

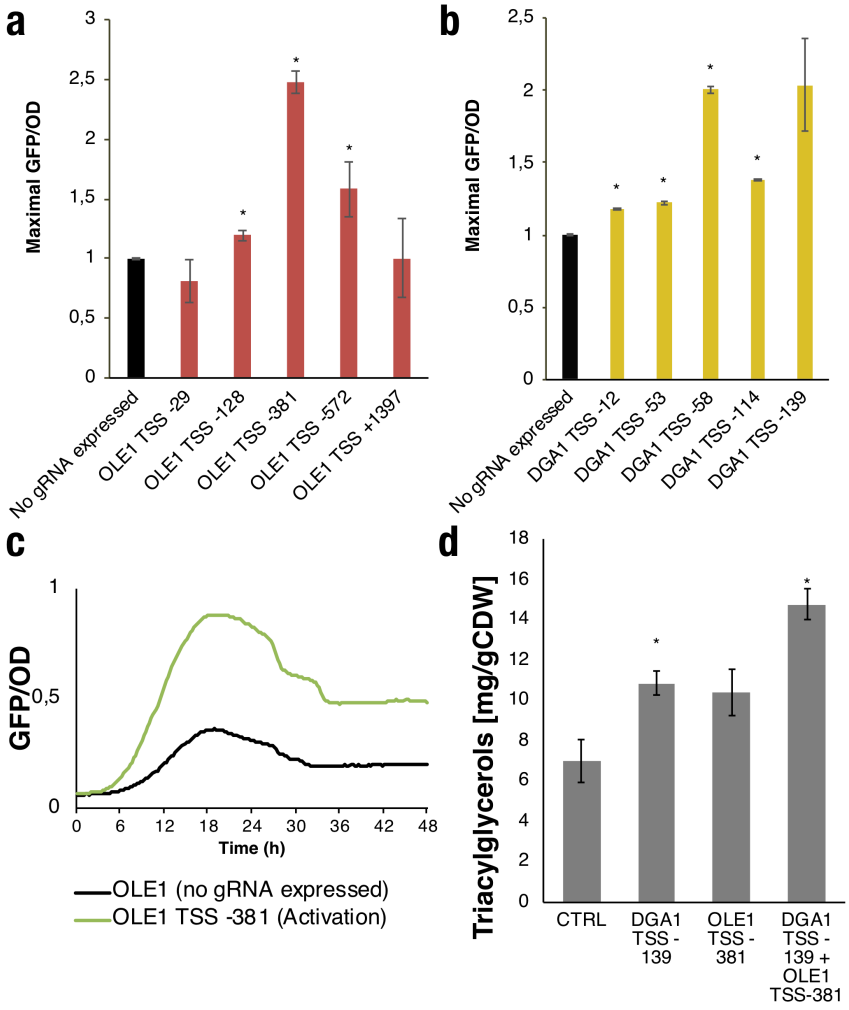


Figure III-3. Transcriptional regulation from gRNAs coupled with dCas9-VPR and their impact on triacylglycerol production. a. Effect of different gRNAs tested for *OLE1* promoter coupled to GFP. **b.** Effect of different gRNAs tested for *DGA1* promoter coupled to GFP. Bar plots for GFP/OD represent expression relative to the control strain without any gRNA expressed. Engineered strains were cultivated with aTc and fluorescence levels were measured using Biolector. The maximum values for GFP/OD yield over the entire cultivation time were then extracted. Displayed are averages ± SD of maximum values of triplicates, * p value < 0.05; (Student's t-test: one-tailed, two-sample equal variance) **c.** Over time monitoring of GFP/OD for OLE1p-GFP with the gRNA OLE1 TSS -381 in green and control (no gRNA expressed) in black. **d.** TAG quantification after 24 h. Strains were cultivated in 2% glucose synthetic minimal media. * p value < 0.05; (Student's t-test: one-tailed, two-sample equal variance).

Here, from a small library of five gRNAs per promoter designed to span over both promoters, gRNAs *DGA1* TSS-139 (i.e. gRNA targeting *DGA1* promoter at position TSS-139, which corresponds to the distance from the TSS to the middle of the gRNA) and TSS-58, and *OLE1* TSS-381 resulted in the strongest upregulation, 2 and 2.5-fold activation, respectively (Fig. III-3). Notably, *OLE1* TSS-29 slightly downregulated the GFP expression (Fig. III-3). Targeting at close vicinity to the TSS might have hindered correct docking of the RNA pol. II, as described in previous reports (161).

Finally, based on this screening, we sought to assess whether the transcriptional effect obtained from the gRNAs leading to highest expression levels could be translated into higher fluxes towards our desired compound. Single expression of *OLE1* TSS-381 and *DGA1* TSS-139 led to a 1.5-fold increase of TAG levels and 2-fold increase when expressed together, compared to control, i.e. no gRNA expressed (Fig. III-3).

3.2 Discovering novel gene regulatory setups using targeted gRNA libraries (Paper V)

In the previous sections, we rationally applied Cas9 and dCas9-based technologies to efficiently rewire metabolic pathways towards the production of relevant compounds. For dCas9, we showed that expression fine-tuning can be obtained by designing a small gRNA library guided to the promoter of interest. While this method has been proven efficient for a small set of promoters, it is challenging to apply similar screenings when targeting a large number of promoters at a genome-wide level. In this context, genetically encoded biosensors, when coupled to fluorescent output, can help enhancing the throughput by enabling real-time monitoring of cellular metabolism and screening of large diversified libraries using FACS (162) (Fig. III-4). Genome-wide approaches are particularly attractive in a metabolic engineering context as they allow to screen, from a population of cells, particular setups, e.g. gene regulation or mutations, favoring a specific phenotype. Conversely to our previous approaches where we focused on specific pathways, these approaches enable the elucidation non-obvious genes linked to the enhancement of the phenotype desired. As such, in **Paper V** we sought to exploit the high-throughput feature of a transcription factor-based biosensor together with the versatility and fine-tuning properties of CRISPR-based transcriptional regulation applied to discover novel gene expression fine-tuning set-ups enhancing fluxes towards precursors for oleochemical production.

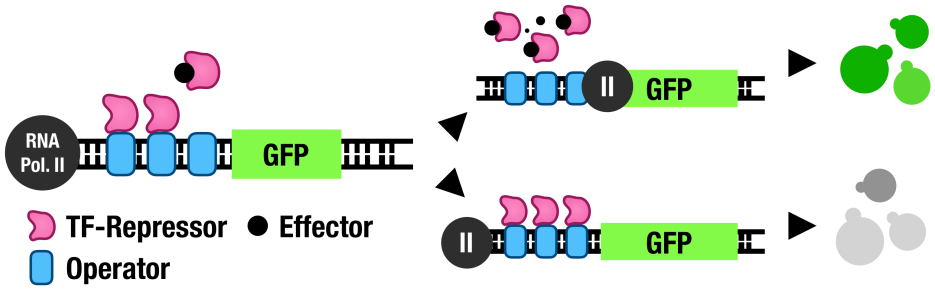


Figure III-4. Transcriptionally regulated biosensor. The transcription factor (repressor) binds to its target sequence (operator) in the absence of the metabolite effector (effector). The GFP expression is modulated based on the abundance of the effector, enabled in high levels and repressed otherwise. This system allows distinguishes high-producer of effectors based on the fluorescence levels and can ultimately be selected using FACS.

From the various characterized biosensors in yeast, we selected the malonyl-CoA responsive transcription factor FapR from *Bacillus subtilis*, previously characterized by David and colleagues (163). FapR binding sites (*fapO*) were integrated into a *TEF1* promoter (*TEF1p*) to control GFP expression as the output signal. FapR is a homodimeric repressor which binds to its operator (*fapO*) in absence of malonyl-CoA and detaches when bound to malonyl-CoA. Malonyl-CoA is an important precursor for various industrially relevant compounds, e.g. fatty-acid derived products as it is, together with acetyl-CoA, the main substrate of fatty-acid biosynthesis and various other products such as the bioplastic precursors, 3-hydroxyprionic acid (3-HP) (164).

Here, while it was theoretically straight-forward to couple a genome-wide gRNA library to a FapR biosensor-based screening, three important questions remained concerning the experimental design:

1. Given that there are thousands of genes whose regulation could potentially affect malonyl-CoA levels, and expression of each of those genes would need be fine-tuned, which genes should be targeted?
2. We previously observed in **Paper IV** that many of the gRNAs covering *OLE1* and *DGA1* promoters did not lead to any substantial change in expression level. Thus, how many gRNAs per promoter need to be included?
3. How can simultaneous down- or up-regulation setups using a single system be achieved?

We used an *in silico* approach based on a yeast GEM to assist on selecting which genes to target (question 1) (165). As briefly introduced in **Chapter I**, GEMs models are mathematical representations of the metabolic network of the cell (40, 166). They are constructed by integrating thousands of reactions and metabolites interacting inside the cell and provide a comprehensive in-depth understanding of the studied organism. One advantage of using GEM is the ability to perform biological predictions at the system level, e.g. predicting the growth rate of the studied organism or the rate of production of a metabolite (167). This is commonly done by Flux balance analysis (FBA) which analyze the flow of metabolites through a metabolic network, assuming a biological objective is applied in the form of maximizing (or minimizing) a certain flux at steady-state growth and mass balance (167). In this study (**Paper V**), the objective function was set to be maximized for cytosolic production of both acetyl-CoA and malonyl-CoA (Fig. III-5).

Once the maximum specific growth rate (μ_{max}) was established, we applied 11 suboptimal growth rates (0.2 to 0.8, increment: 0.05), and acetyl-CoA and malonyl-CoA production were maximized using FBA (Fig. III-7, **Paper V:** Fig. 1a-d, Table S1-2). Here, for each simulation, a *k* score that compared the flux of each reaction with the flux of the same reaction at maximum growth rate conditions was determined. These scores were defined as larger than 1 representing an increased flux compared to the μ_{max} condition, whereas *k* scores lower than 1 show a decrease in the flux. In terms of gene expression, it was

arbitrarily defined that *k* score represented upregulation or downregulation of the gene expression levels and genes with a *k* scores between 0.549 and 1.001 were ruled out.

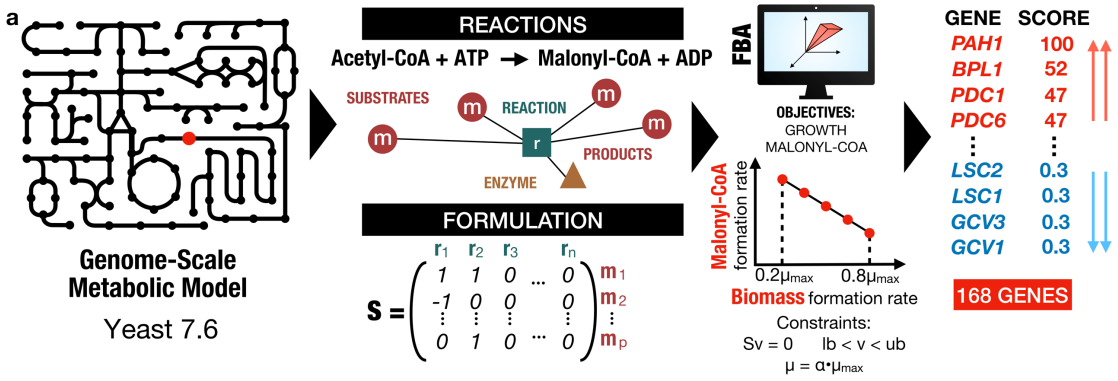


Figure III-5. Flux-balance analysis to retrieve genes potentially enhancing malonyl-CoA production. Genome-scale metabolic model fundamentals where each metabolic reaction and its constituents are mathematically represented in a global model. Here, we applied FBA maximizing the production of malonyl-CoA as the objective function of the model. Fluxes towards malonyl-CoA are computationally evaluated at different growth rates and a list of genes potentially contributing to enhancing malonyl-CoA production is retrieved. A target score for upregulation or downregulation is associated to each gene as a proxy for fine-tuning. Here, for all α's and all reactions, the scoring is defined as the average of the $flux_{\alpha}$ over $flux_{WT}$

These *in silico* analyses ultimately scored genes that are expected to affect fluxes towards the production of acetyl-CoA and malonyl-CoA at different growth rates, and ultimately reduced the number of genes to target to only 168 genes, with 70 genes targeted for upregulation, 80 for downregulation, and 18 found to be either upregulation or downregulation targets depending on the carbon source (**Paper V**: Table S2) (Fig. III-5).

To answer the second and third questions, we described in the previous section how the constitutive CRISPR-dCas9 system could potentially enable both activation and downregulation with the help of RNA-stem loops fused to the gRNA. However, we observed that this system was underperforming compared to the inducible system. Furthermore, gRNAs targeting the *OLE1* promoter in the close vicinity to the TSS led to a downregulating effect even when guiding dCas9-VPR (*OLE1* TSS -29) (Fig. III-3). Thus, to both counteract the uncertainties that influence gRNA regulatory efficiency and reach optimal transcriptional interference (up- or downregulation), we designed up to 21 gRNAs covering each targeted gene, including up to five adjacent to their respective TSS. This design yielded to a library of 3194 gRNAs (**Paper V**: Table S2).

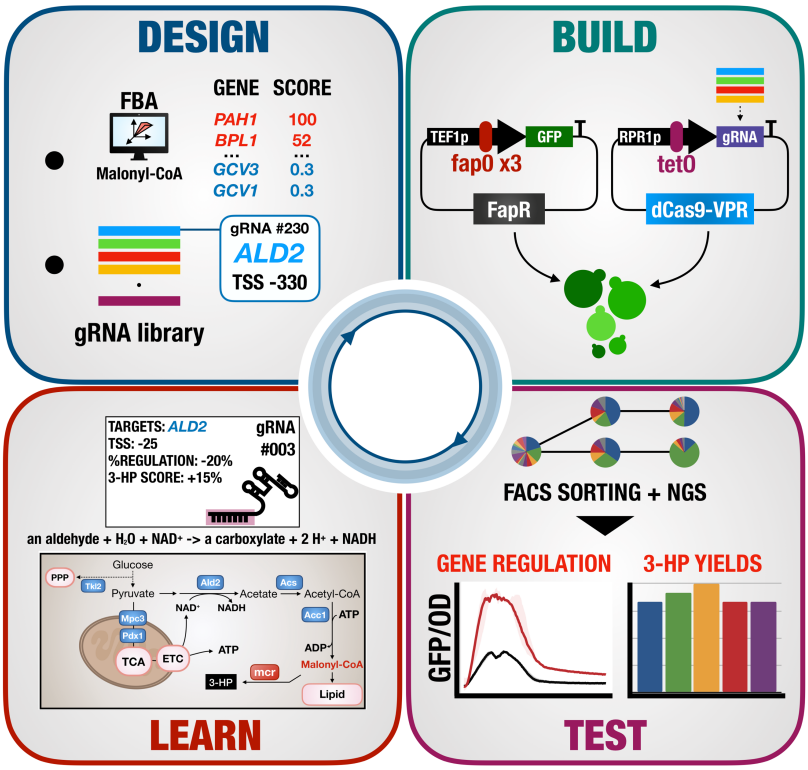


Figure III-6. DBTL cycle applied to the study. Design. The gene scores are obtained by FBA, a gRNA library of 3194 gRNAs is designed to target these genes. **Build.** The gRNA library is cloned into a dCas9-VPR plasmid and transformed into a yeast strain expressing the FapR-based malonyl-CoA biosensor. **Test.** Multiple rounds of fluorescence-based sorting and NGS are applied to retrieve the promising gRNAs, i.e. gRNAs potentially enhancing the flux towards malonyl-CoA. Then, the enriched gRNAs are assessed on their efficiency to improve 3-HP production as well as transcriptional regulation by coupling their targeted promoter to *GFP*. **Learn.** Based on the 3-HP and transcriptional regulation data, we can elucidate the effect of the selected gRNAs on yeast metabolism.

The gRNA library was subsequently cloned into the inducible dCas9-VPR plasmid and inserted into yeast expressing both FapR constitutively and *ACC1*** under control of the *HXT7* promoter (CEN.PK-11C + HXT7p-*ACC1*** + FapR-based malonyl-CoA biosensor). *ACC1*** was chosen in order to channel potential increases of the acetyl-CoA pool towards the malonyl-CoA pool, and thus activating the malonyl-CoA biosensor response. The yeast library was consecutively sorted using FACS over three days using two different fluorescence gates (Fig. III-7). Next-generation sequencing (NGS) was performed at each stage and the gRNA distribution in the population was determined. From this analysis, we subsequently selected the gRNAs that were most significantly enriched throughout the different sorting steps, and ultimately, assess their performance on both their transcriptional regulation, i.e. fine-tuning properties on their respective targeted promoters coupled to GFP, and enhancing fluxes towards malonyl-CoA pools, through quantification of malonyl-CoA-derived product titers, in form of 3-HP (Fig. III-6).

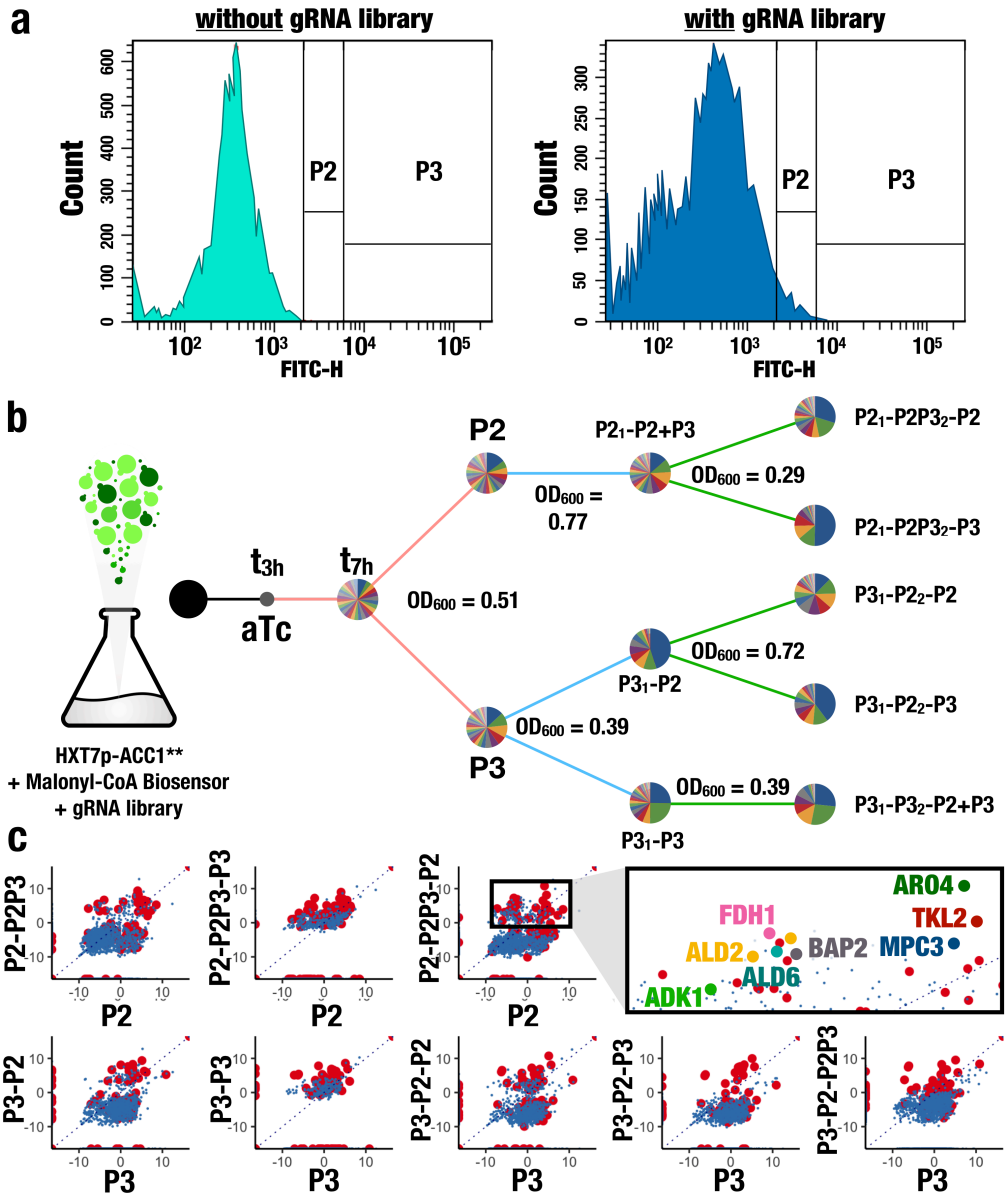


Figure III-7. Three-day sorting strategy of the yeast library. a. After 7 h of culture, the yeast library was subsequently sorted using two different gates, one for medium fluorescence, P2, and one for high fluorescence, P3. The yeast strains without the library (empty plasmid) in the left show few cells on P2 and P3. **b.** Three-day sorting using the two gates, except for $P2_1$-P2+P3 where $P2_1$-P2 and $P2_1$-P3 were accidentally mixed together. Each pie chart represents the gRNA distribution after the sorting. The colored lines represent the day of sorting: black for day 0; red for day 1; blue for day 2; and green for day 3. **c.** Log2 fold-change gRNAs over the initial library (pre-culture). Shown in red are the most enriched gRNAs that were subsequently tested for 3-HP.

From the different sorting schemes, we selected 49 gRNAs, targeting a total of 46 genes, that were enriched after three rounds We subsequently validated their efficacy to enhance malonyl-CoA levels by evaluating their expression on the production of 3-HP. It can be obtained by expressing the bifunctional malonyl-CoA reductase Mcr from *Chloroflexus aurantiacus* which directly catalyzes the conversion from malonyl-CoA to 3-HP in a two-

step reaction (168). Expression of Mcr formed a direct metabolic pull likely converting any (even temporary) increase in malonyl-CoA into 3-HP accumulation over time. Here, among all gRNAs selected, only a few ultimately increased the overall 3-HP yield by more than 10% (Fig. III-8). We selected 14 of these genes and coupled their respective promoters to GFP expression (Fig. III-8, **Paper V**: Fig. S4, Table S5).

Analysis of the interesting gRNAs

We observed that most efficient gRNAs target genes whose corresponding enzymes use NAD^+ as co-factor, e.g. cytoplasmic aldehyde dehydrogenase Ald2*,* glyceraldehyde-3-phosphate dehydrogenase Tdh2, or, expectedly, genes located upstream of malonyl-CoA pathway, e.g. Acs2, Ald2*,* Pdx1 (Fig. III-8; **Paper V**: Fig 2c, Table S4). The highest increase in 3-HP yield (36% increase) was achieved by a gRNA targeting the gene encoding adenylate kinase 1, *ADK1* #15 (TSS -761), which also showed a downregulating effect on the activity of the promoter coupled to GFP (Fig. III-8). Here, if regulating *ADK1* causally impacts 3-HP, I speculate that this regulation might have led to *(i)* a decreased availability of ADP, an important metabolite that positively regulates Acc1 regulator Snf1, and *(ii)* an increased level of ATP, which is used in the biosynthesis of malonyl-CoA from acetyl-CoA (169–171). Interestingly, the gRNA also binds downstream of *HTA1*, which encodes one of the main histone proteins involved in the chromatin structure. This binding might have reduced *HTA1* mRNA levels by blocking RNA polymerase II and thus potentially changed chromatin structures and consequently the expression levels of other genes. Targeting *ALD2* gRNA #1 (TSS -241) increased 3-HP yield by 27% compared to the control, and also showed a significant increase in GFP (Fig. III-8). This gene is repressed in the glucose phase (172), and we speculate that in a context where NAD^+-dependent enzymes seem to be favored for increased fluxes towards malonyl-CoA, de-repressing it, eventually benefited 3-HP production. *ARO4* #18 (TSS -641) expression led to a 22% increase in 3-HP yields but, surprisingly, did not influence *ARO4* transcriptional levels (Fig. III-8). As with *ADK1* #15, we noticed that *ARO4* #18 (TSS -641) binds closer to the adjacent gene, *SPO23* (*YBR250W*). As such, the effect on the *SPO23* promoter coupled to GFP led to a significant upregulation (Fig. III-8). *SPO23* has not yet been characterized which makes it challenging to speculate on the underlying reasons behind the increase in 3-HP yields. However, Tevzadze and colleagues highlighted a close interaction between Spo23 and Spo1, a meiotic PI-specific phospholipase B, which could indicate a link with fatty acid turnover (173).

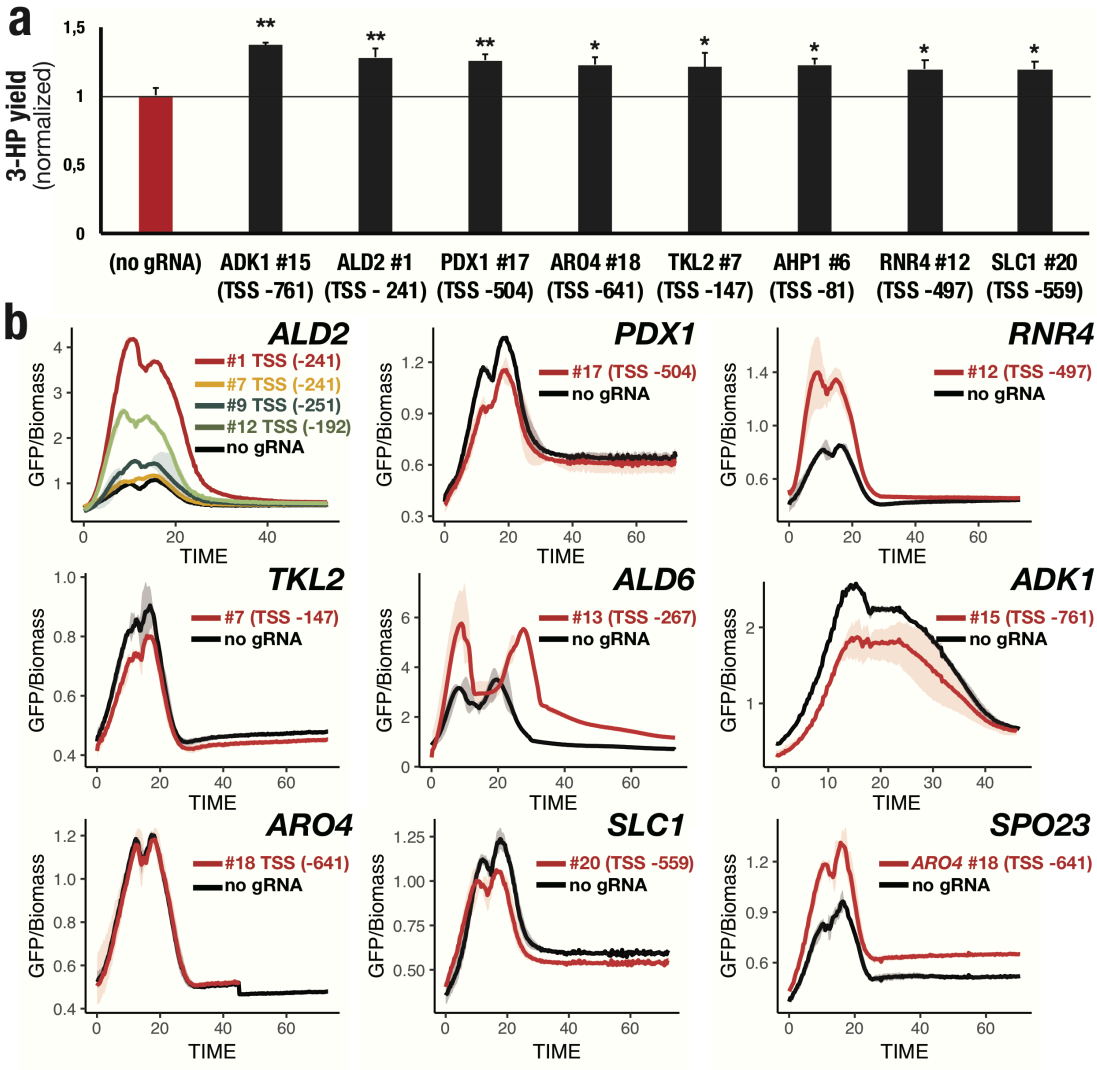


Figure III-8. Effect of the enriched gRNAs on enhancing fluxes towards malonyl-CoA. a. Shown are the yields increased by more than 15% the total 3-HP levels normalized to OD and control (no gRNA expressed). Values were obtained in triplicates grown in defined minimal medium with 20 g·L^{-1} glucose and cultures were sampled after 72 h. Displayed are averages ± SD. ** p value < 0.01; * p value < 0.05; (Student's t-test: one-tailed, two-sample equal variance). d. Influence of different gRNAs on the activity of the respective target promoter. Shown are TKL2 #7 (TSS -147); and ALD2 #1 (TSS - 241), #7 (TSS -241), #9 (TSS -251), and #12 (TSS -192). **b.** Mean fluorescence intensities (GFP/OD) were obtained from three biological replicates ± S.D. monitored with a BioLector.

Finally, we also identified new candidate genes to increase 3-HP production such as genes involved in the processing of precursors upstream of malonyl-CoA, e.g. *TKL2* encoding transketolase, and *PDX1* encoding a subunit of the pyruvate decarboxylase complex as well as *AHP1* encoding a peroxiredoxin that protects against oxidative stress (Fig. III-8). Notably, *PDX1* #17 (TSS-504) shows a downregulating effect on *PDX1* promoter and this effect might have enhanced the flux cytosolic conversion of pyruvate to acetyl-CoA instead of the NAD^+ consuming mitochondrial one (Fig. III-8) (174).

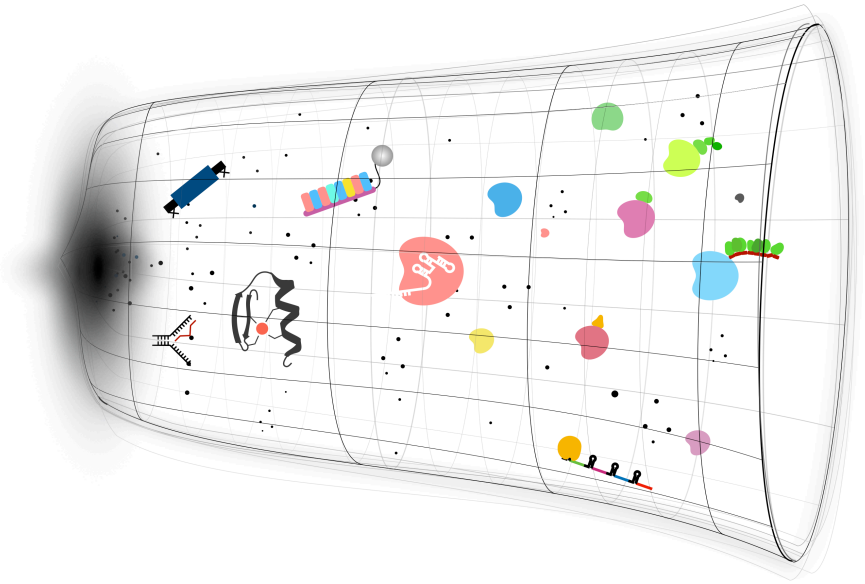
3.3 Conclusions of Chapter III

In **Chapter III**, we explored dCas9-based technologies for fine-tuning of biosynthetic pathways. As previously mentioned, achieving correct expression of the genes involved in the studied pathway as well as determining which genes to reprogram remain currently challenging to achieve with traditional engineering strategies. In **Paper IV**, based on multiple gRNAs tested, we established dCas9-VPR as a robust tool for rewiring metabolic fluxes in yeast, among the different CRISPR-dCas9 tools. This was concluded by, first, monitoring the fluorescence levels of the targeted promoter to GFP upon the gRNA expression, and by quantifying TAG levels (as these gRNAs were targeting genes involved in lipid biogenesis pathways).

Then, in **Paper V,** we sought to apply the same methodology at a larger scale. We established a framework based on GEM simulations for predicting key genes that potentially favor the production of the endogenous metabolites and exploited the versatility of dCas9-VPR for fine-tuning gene expression coupled with an intracellular biosensor for selecting high-producing cells. Here, our *in silico* constraint-based modeling approach predicted a specific set of 168 genes whose regulation potentially impact the production of malonyl-CoA. Then, combining a gRNA library targeting these genes and a malonyl-CoA responsive intracellular biosensor in yeast ultimately enabled to retrieve optimal gene expression of both known and novel genes contributing to enhancing the flux towards cytosolic malonyl-CoA. However, while we observed that 3-HP production was improved upon fine-tuning genes involved in providing malonyl-CoA precursor, co-factor supply, as well as chromatin remodeling, many of the selected gRNAs were not influencing 3-HP levels. This indicates that either these gRNAs might still increase malonyl-CoA levels but not sufficiently to be converted to 3-HP, or that the selection parameters I chose for the study were not optimal. Another possibility could come from the discrepancy between the fluorescence sorting performed approximately at the glucose phase and the 3-HP sampling, performed at 72 h. Thus, enriched gRNAs that marginally enhance malonyl-CoA during the glucose phase might not be enough to reach a significant increase in 3-HP yields.

While applied for malonyl-CoA, the same approach can be could in theory be applied for other metabolic engineering studies. Thus, I developed an R shiny app that allows users to easily choose a metabolite to optimize in yeast. Once the metabolite selected, the tool generates a list of suggested genes to regulate with a detailed list of gRNA targeting them (hosted at https://raphdl.shinyapps.io/crispri-gem/).

Chapter IV: Expanding the CRISPR repertoire - Developing new tools to accelerate strain engineering



4.1 Increasing the targeting range through multiplexing (Paper VI)

When conceptualizing **Paper VI** in early 2016, the research landscape that applied multiplexing features for yeast strain design remained at its infancy (175, 176). Thus, we aspired to contribute to the yeast scientific community by developing a multiplexing tool for both genome engineering and gene regulation.

CRISPR-associated endonuclease Csy4.

As described in **Chapter I**, the type I CRISPR-system is defined by having the crRNA guiding a cascade complex made of several proteins to target invading viral DNA (Box 1). The interference is then performed by the Cas3 subunit. In *Pseudomonas aeruginosa*, the cascade complex is composed of Csy enzymes, with Csy4 binding and processing the pre-crRNA in crRNAs (Fig. IV-1). This endoribonuclease has been shown to have a high degree of substrate binding towards a 28 nucleotide RNA stem-loop (5'-GUUCACTGCCGTATAGGCA**G**CUAAGAAA-3') and cleaves after the guanine at position 20 downstream (underlined in bold) (Haurwitz et al. 2010, Haurwitz et al. 2012). Additionally, previous structural studies have revealed the ability of Csy4 to discriminate its crRNA from other cellular RNAs. Altogether, Csy4 is an ideal candidate for attempting to carry out multiplexing experiments (Fig. IV-1).

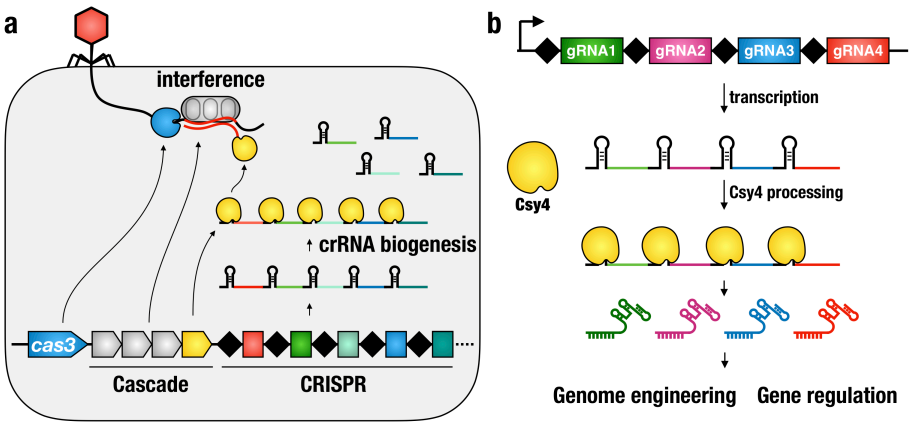


Figure IV-1. Utilization of elements from Type I CRISPR-systems for genome editing and gene regulation. a. Overview of Type I CRISPR-system. The *csy* complex is usually composed of several *csy* genes, including *csy4* (in yellow). Csy4 processes the CRISPR array to generate several crRNA which will then be used for the interference step against phages. **b.** Harnessing Type I Csy4 from *P. aeruginosa* for gRNA multiplexing.

Genome editing and gene interference applications in S. cerevisiae.

We first sought to characterize the ability to perform multiple gene knockouts by simultaneously generating gRNAs from a single transcript using Csy4 in yeast. Thus, starting from a strain constitutively expressing Cas9 (IMX581: *MATa MAL2-8c SUC2 ura3-52 can1::PTEF1-cas9*), we constructed and genomically integrated the gene encoding Csy4 (with a nuclear localization signal) expressed under control of *TEF1p*.

Next, we designed an array of four gRNAs in a row, targeting genes involved in the fatty acid metabolism: *FAA1*, *FAA4*, *POX1*, and *TES1*, interspaced by the Csy4 28nt recognition sequence (Fig. IV-2). These gRNAs were previously evaluated in other projects (e.g. **Paper II**) thus their high-efficiency was known. The gRNAs transcript was expressed under the constitutive RNA Pol. III promoter, *SNR52p*, from a multi-copy 2µ plasmid. The plasmid was transformed along with linear repair DNA fragments matching the flanking regions of the targeted genes, causing a deletion of the genes. Expression of Csy4 yielded close to 100% efficiency of quadruple knockouts, whereas in absence of Csy4 only double knockouts were detected. These results confirmed the ability to use Csy4 for multiplexed knockouts (Fig. IV-2). Notably, the reasons behind achieving a double knockout even in absence of Csy4 could potentially be the instability of the scaffold due to the many repetitive elements (28nt + 79nt gRNA scaffold) coupled with the fact that the gRNAs being expressed in a multi-copy plasmid and under a constitutive promoter might have generated partial transcripts.

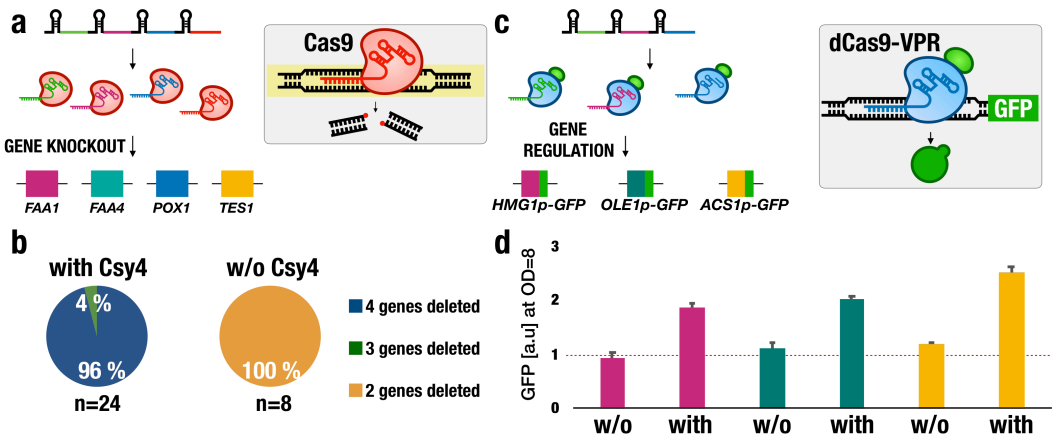


Figure IV-2. Multiplexing genome editing using Csy4. a. Multiplexing of four gRNAs targeting *FAA1*, *FAA4*, *POX1*, and *TES1*, for gene knockouts. **b.** Expression of Csy4 allows near-perfect quadruple knockout (in 24 colonies tested). **c.** Multiplexing of three gRNAs targeting *HMG1p*, *OLE1p*, and *ACS1p*, for gene regulation using dCas9-VPR. **d.** Fluorescence quantification at OD_{600} = 8. Red dotted line defines levels of GFP when no gRNA is expressed. Pink: *HMG1p-GFP*; Turquoise: *OLE1p-GFP*; Yellow: *ACS1p-GFP*

We next sought to assess the capacity of our platform for efficient transcriptional regulation applications. As with the knockout experiment, we designed an array of three gRNAs, which have been previously characterized (**Paper IV**), interspaced by the 28 nt sequence and expressed them from a single copy plasmid under control of a tetracycline-inducible RPR1 promoter (70, 157). The multiplexed and Csy4-processed gRNAs reached the same level of transcriptional regulation as to when the particular gRNA was expressed alone P_{OLE1} and P_{HMG1} (Fig. IV-2, Fig. IV-3). For *ACS1*, the processed gRNA-transcript reached almost the same level as the positive control in the presence of Csy4 (Fig. IV-3). Here, we notice that the in absence of Csy4, GFP levels were similar to the control (empty plasmid). This indicates the importance of controlling the expression of gRNAs to produce significant interference, i.e. using a single-copy plasmid might prevent the generation of partial transcripts interfering with the experiment, as we saw in with the genome engineering experiments.

In conclusion, we showed that Csy4 specific endoribonuclease activity can efficiently be utilized for multiplexing of CRISPR-mediated gene deletion and interference in *S. cerevisiae*.

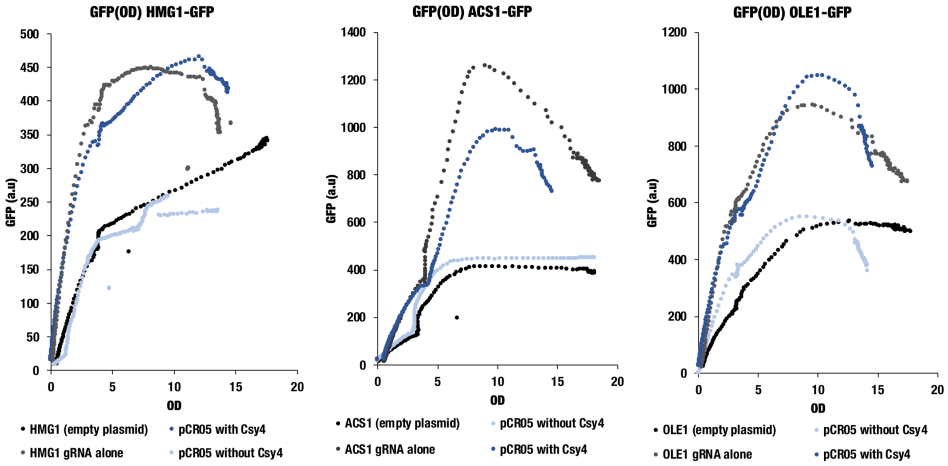


Figure IV-3 Detailed transcriptional regulation with and without Csy4. Shown here are the GFP in function of OD for one array processed (or not) by Csy4. In black: empty plasmid; grey: unique gRNA; light blue: array of gRNAs without *csy4* expression; blue: gRNA array processed by Csy4.

4.2 Exploiting promiscuous on-target specificity (Paper VII)

In the previous section, we showed that processing an array of gRNAs in a row allowed to target multiple targets at once. Metabolic engineering of microbial cell factories usually requires targeting genes within the same pathway. One example is targeting multiple genes encoding enzymes carrying the same catalytic function, so-called paralogs, e.g. *FAA1* and *FAA4* for preventing FFA degradation, or *TGL3*, *TGL4* and *TGL5* for TAG degradation (**Paper II**, **Paper III**). These paralogs are usually derived from genome or gene duplication and thus show remarkable sequence similarities. Based on this feature, we sought to exploit that feature to accelerate strain design, i.e. exploiting promiscuous on-target idiosyncrasies of certain gRNAs for multiplexed engineering. These gRNAs, that we coined "promiscuous gRNAs", eventually would allow minimizing the use of gRNAs in applications where editing multiple genes simultaneously is desirable (see previous sections).

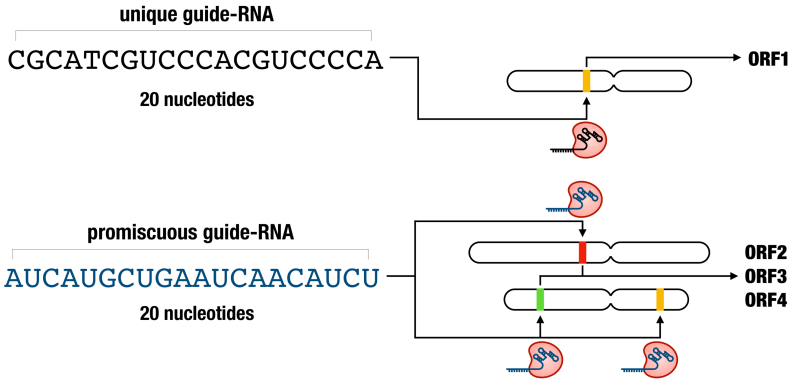


Figure IV-4. Graphical representation of a promiscuous gRNA.

We designed a computational tool to retrieve and explore promiscuous gRNAs. The tool was written in R and is made available on GitHub (https://github.com/gattofrancesco/pgRNA). The algorithm was built to use a genome sequence or a list of genes as input and extract all gRNAs by locating the PAM sequence in the target region (Fig. IV-5). An option to filter poor quality gRNAs, such as gRNAs with repetitive nucleotides, is possible by imposing minimal %GC content. Each gRNA is then aligned back to all the original targets, and any gRNA with several matches is defined as promiscuous gRNA. The probability for a gRNA to find new matches can be increased by allowing mismatches in the sequence. Then, the hits are analyzed, and a list of the most promising promiscuous gRNAs is generated (Fig. IV-5).

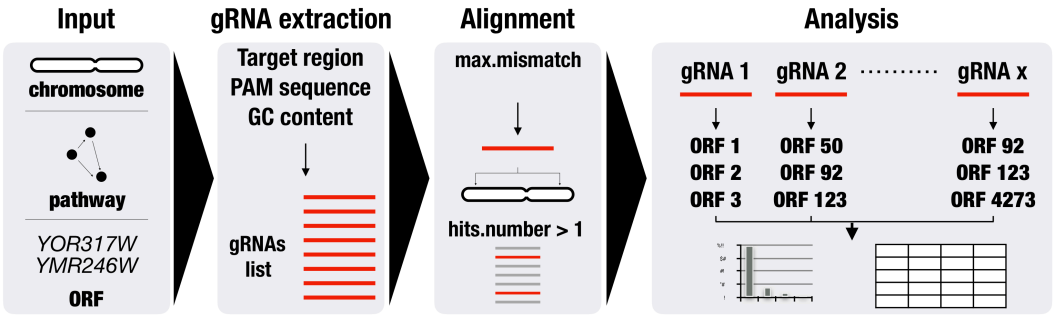
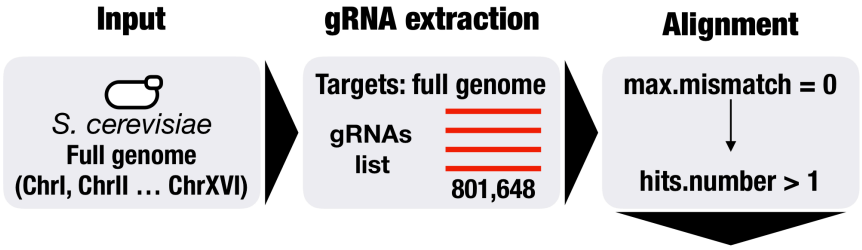


Figure IV-5. Workflow of the promiscuous gRNA tool.

We chose to apply our tool to identify promiscuous gRNAs targeting multiple sites in the *S. cerevisiae* genome and experimentally validate some of them (Fig. IV-6, Fig. IV-7). For the whole genome, 801,648 individual gRNAs were retrieved for all sixteen nuclear chromosome sequences, with the most promiscuous gRNA ("GCAAGGATTGATAATGTAAT") binding a total of 148 times. The promising promiscuous gRNAs expectedly showed a significant number of hits in repetitive elements such as Ty retrotransposons and tRNA regions. While I have unsuccessfully tried to experimentally validate these highly promiscuous gRNAs with the aim of removing these repetitive elements, in **Paper VII** we exclusively targeted coding sequences.



	chrI	chrII	chrIII	chrIV	chrV	chrVI	chrVII	chrVIII	chrIX	chrX	chrXI	chrXII	chrXIII	chrXIV	chrXV	chrXVI	total hits
GCAAGGATTGATAATGTAAT	2	7	8	25	9	3	15	5	2	11	4	15	12	8	10	12	148
TATACTAGAAGTTCTCCTCG	2	3	9	14	6	2	14	5	1	7	1	11	8	9	8	9	109
TACTAGTATATTATCATATAC	3	3	6	16	8	2	6	2	1	4	0	11	4	5	5	6	82
GAAATAGTCATCTAAATTAG	2	3	6	16	7	2	6	2	1	5	1	11	4	5	5	6	82

Figure IV-6. Promiscuous gRNAs for *S. cerevisiae*'s whole genome.

We next set out to experimentally validate several predicted promiscuous gRNAs. From the initial extracted gRNAs, only 1% had two or more targets. From this list, we limited our search to promiscuous gRNAs targeting 2 genes in the lipid metabolism and tested their efficacy via gene deletion (Fig. IV-7). We selected in a first experiment, a promiscuous gRNA targeting *FAA1* and FAA4, and in a second experiment, a promiscuous gRNA *PLB1* and *PLB2* (Fig. IV-7). Finally, correct ratio of successful deletion reached 100% for both *FAA1/4* and *PLB1/2* experiments (Fig. IV-7). In conclusion, this study proposes a computational approach to exploit promiscuous on-targeting effects of gRNAs to consent multiplexed genome editing in a time and cost-effective fashion.

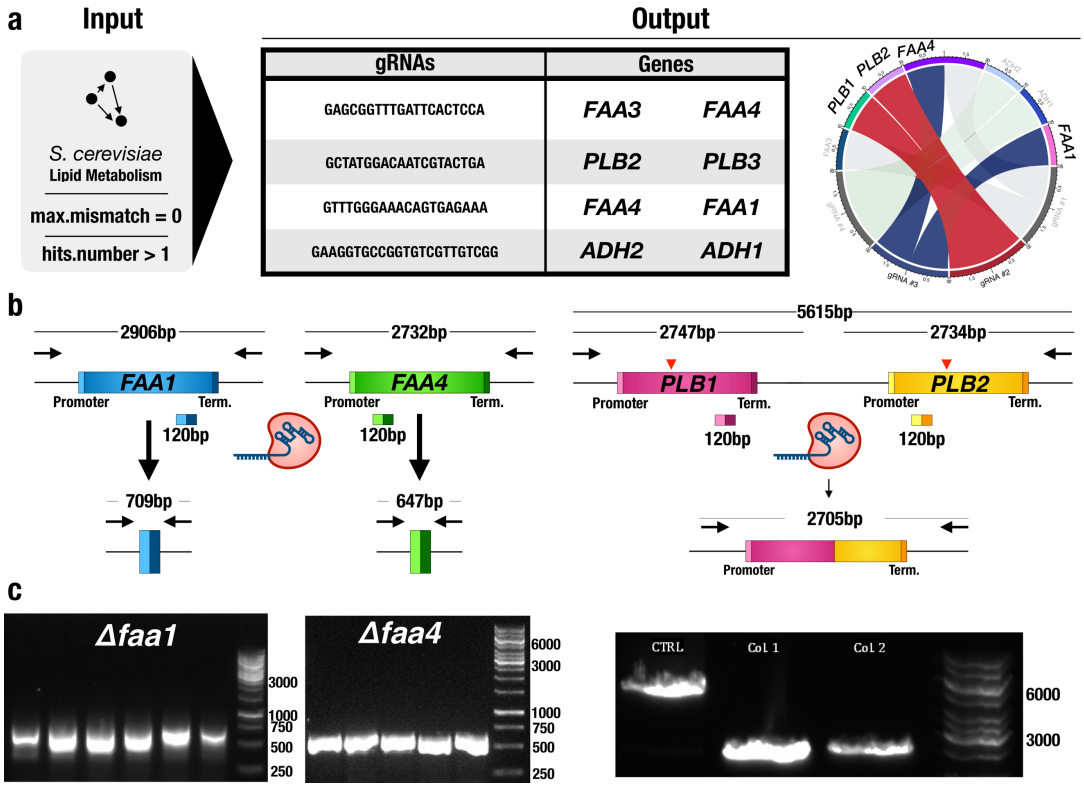


Figure IV-7. Application of promiscuous gRNAs for genome editing in *S. cerevisiae*. a. Computational output of the algorithm when taking lipid metabolic genes as input. **b.** Two promiscuous gRNAs tested, one targeting both *FAA1* and *FAA4*, the other targeting both *PLB1* and *PLB2*. Along with the promiscuous gRNA, the strains were co-transformed with repair fragments matching the flanking regions of the genes targeted. **c.** Agarose gel showing colony PCRs. Correct deletions for *FAA1* and *FAA4* showed bands at 709 and 647 bp, respectively. For *PLB1* and *PLB2*, correct deletion showed a band 2705bp obtained by NHEJ, while the control (CTRL) shows a band at 5615bp.

4.3 Cas12 and Cas13

Besides Csy4, I, together with colleagues, have attempted to characterize other types of CRISPR-systems in yeast, namely Cas12 and Cas13 (Fig. IV-8).

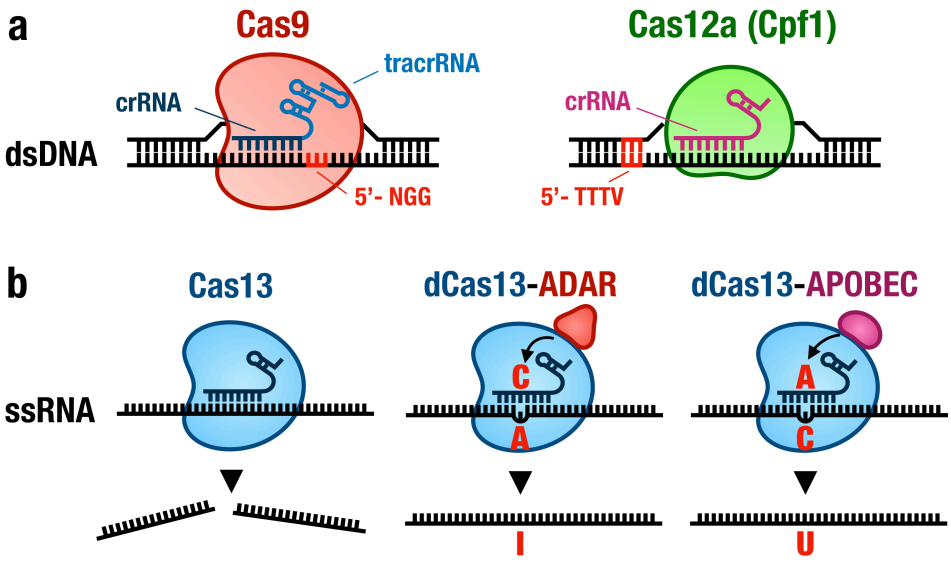


Figure IV-8. Cas12a (Cpf1) and Cas13 features. a. b. Cas13 performs targeted cleavage of an ssRNA substrate. dCas13-ADAR RNA base editing converts A to I (read as G). dCas13-APOBEC RNA base editing converts C to U.

Cas12a

Cas12a, initially known as Cpf1, is an ortholog of Cas9 that offers several attractive features: *(i)* it possesses an RNA processing domain that enables CRISPR array processing into multiple crRNAs (74); *(ii)* its resulting DNA cleavage leaves a 5′ overhang instead of a blunt-end cleavage for Cas9 (73); *(iii)* it has target specificity superior to Cas9 (177); *(iv)* it recognizes a T-rich PAM; and *(v)* it is slightly smaller than Cas9, thus more attractive for clinical applications (73) (Fig IV-8). Following the same approach as the one employed for Csy4, I aimed to apply this system for genome engineering and gene regulation in *S. cerevisiae*. At the time of conceptualizing the project, the Cas12a from *Francisca novocida* (FnCpf1) had been described to cut plasmid DNA integration in *E. coli* (73), and further characterization of its dual-nuclease (RNA and DNA) characteristic had been described from Charpentier's lab (74). Based on these two studies, I aimed to characterize the ability to generate multiplexed gene knockouts by attempting to remove *FAA1* and *FAA4,* as well as gene activation by targeting *OLE1*. While I have been successful at removing *FAA1*, multiplexed removal of both *FAA1* and *FAA4* turned out to be more problematic (Fig. IV-9). Also, my attempts to characterize a dead-Cpf1-VPR did not yield any transcriptional regulation when targeting the *OLE1* promoter (data not shown).

Finally, over the course of my efforts, Świat and colleagues successfully characterized FnCpf1 for multiplexed gene editing in yeast (up to four gene deletions) (178), and more recently, Liao and colleagues applied dCas12 for gene activation (179).

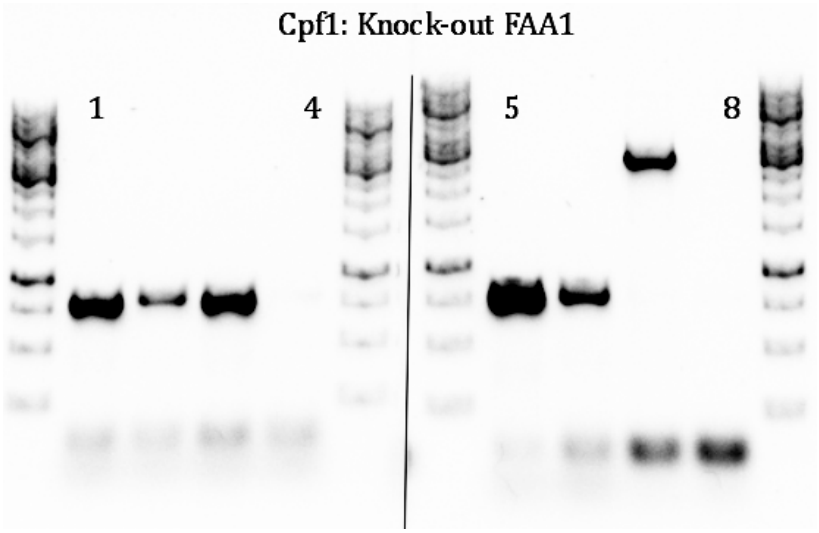


Figure IV-9. Cas12a (Cpf1) proof-of-concept for gene knockout. Colony PCRs for the deletion of *FAA1*. Bands shown at 709bp represents correct deletion, while the band at 2900bp for colony 7 shows wild-type genotype.

Cas13b

Cas13 systems, initially known as C2c2, have recently been developed for precise transcriptome engineering (180, 181). The Cas13 RNA targeting nuclease from *Prevotella sp.* (Cas13b) is a single-effector nuclease specialized against RNA phages (Fig. IV-8) (182). A catalytically inactive mutant (dCas13) has, in the recent years, been developed for RNA single-base editing via fusing it to deaminase enzymes (183). In general, deaminases irreversibly catalyze the conversion of the nucleotide NH_2 group to a ketone leading to a base conversion. The only published example of adenine deaminase applied in RNA editing technologies is ADAR2 which converts adenine to inosine (A to I). Inosine is read as guanine during translation. For correct base-editing, a mismatch in the gRNA is required to selectively deaminate the single mismatched nucleotide. At the time of conceptualizing this project, RNA base editing was limited to adenine to inosine transitions, but has since July 2019 been extended to cytidine to uridine conversion (C to U) (184). Thus, we sought to characterize ADAR2 A to I RNA editing in yeast. Additionally, we attempted to develop a C to U RNA editing system by fusing dCas13b to the cytidine deaminase APOBEC1 (Fig. III-8). To validate both systems we generated two mutated forms of GFP with an early stop codon aimed to be reverted to the original codon by either dCas13-ADAR through an A to I edit, or dCas13-APOBEC through C to U (Fig. IV-10). Unfortunately, while we tested several gRNAs (seven) with various spacer lengths for each system, none of them were able to restore GFP (Fig. IV-10).

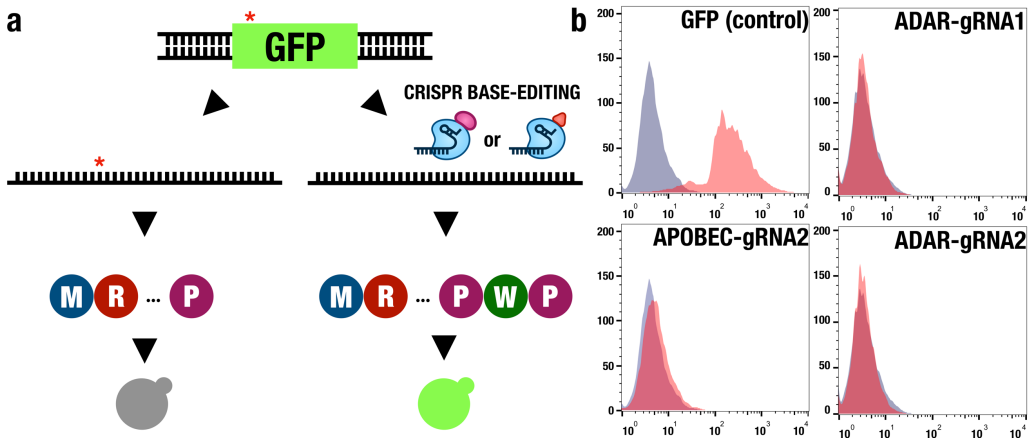


Figure IV-10. System for testing base edits of dCas13-ADAR and APOBEC. a. The red asterisk in GFP represents the early stop codon mutation aimed to be reverted to the original codon by either ADAR through an A to I edit, or APOBEC through an C to U edit. **b.** Flow-cytometry of cells expressing the gRNAs tested. In gRNA histograms, blue represents fluorescence with an empty plasmid and red with the gRNA expressed. For the control histogram, red is the positive control, i.e. unmutated GFP expression.

Notably, during our work in the lab, Jing and colleagues reported the successful implementation of an ADAR base editor in *Schizosaccharomyces pombe*, using a dCas13 from *Leptotrichia wadei* (dCas13a-ADAR), with an editing efficiency reaching up to six percent (185). Here, one can speculate that base-editing in our systems possibly existed but not to the extent to be significantly observed in our flow-cytometer experiment. Here, NGS on the targeted sequence could have provide more insight to this question. Therefore, if base-editing occurs using in our system, it requires further optimization.

4.4 Conclusions of Chapter IV

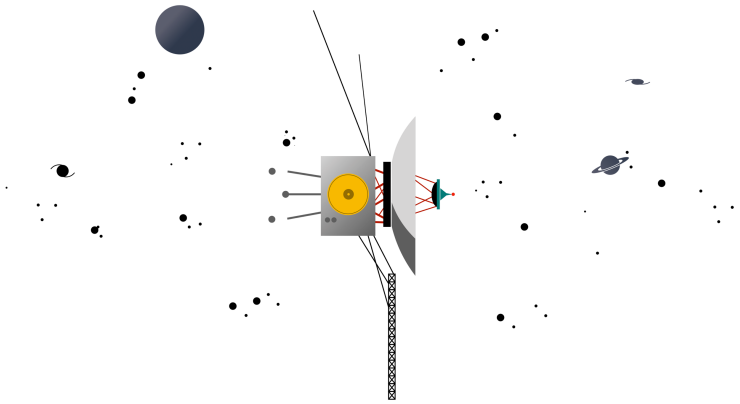
In this **chapter**, I sought to highlight the diversity and versatility offered by CRISPR technologies. The constant invasions of phages and mobile genetic elements has led to the evolution of prokaryotes and archaea with the adaptive and heritable CRISPR system, ultimately yielding a plethora of diverse elements that can be exploited for genetic engineering (186). For example, in **Paper VI**, we harnessed the Type-I Csy4 from *P. aeruginosa* for multiplexed genome engineering and transcriptional regulation *via* processing an RNA transcript into multiple gRNAs. For genome editing, we performed a quadruple deletion with near-perfect efficiency out of 24 colonies tested, and for gene regulation, we showed that a Csy4-processed array of three gRNAs could reach similar transcriptional regulation levels as individually expressed gRNAs. Csy4 has not been, to date, extensively applied in yeast. Possible reasons might be due to the existence of self-cleaving elements, which removes the need of heterologous genes for multiplexing. However, Ferry and colleagues, for example, constructed a switchable gRNA by connecting the crRNA sequence with the 28 nt Csy4 target sequence, ultimately locking the gRNA and preventing it from binding to the genome without Csy4 processing (187).

Also, McCarthy and colleagues recently expanded our work (**Paper VI**) by achieving multiplexing of 12 gRNAs under an RNA pol. II promoter, drastically facilitating the cloning procedure (188).

Additionally, we tried to characterize Cas9 orthologs in yeast, namely Cas12a and Cas13b, for genome engineering and gene regulation, and transcriptome engineering, respectively. While Cas12a showed promising results in terms of gene deletion, I have failed to achieve both array processing of two gRNAs in a row ($gRNA_{FAA1}$ followed by $gRNA_{FAA4}$) and transcriptional regulation (Fig. IV-9). Notably, the direct repeats I selected for *FAA1/4* were 36 nt long, which Świat and colleagues have reported lower efficiency of RNA processing by Cas12a in comparison to 19 nt repeats (19-37% vs. 100%) (178). This result, alongside with selecting an intrinsically poorly efficient gRNA, could have been the reasons of the unsuccessful multiplexing (178). For base-editing using dCas13 CRISPR-systems, based on the study on *S. pombe* from Jing and dCas13a-ADAR where they achieved up to six percent success rate of base-editing, future approaches could take into consideration focusing on dCas13a instead of dCas13b (185). Additionally, restoring GFP via base-editing might not be the appropriate system when the rate of base-editing is extremely low. Here, NGS would give a better readout, as the expected base replacement would be quantifiable by analyzing the reads.

Most of the currently available bioinformatics tools for designing gRNAs only suggest gRNAs uniquely targeting in the host genome, i.e. without any off-targets. In **Paper VII**, we designed a bioinformatic tool opposite to this paradigm, i.e. a tool exploiting the off-targeting of gRNAs for accelerating genome engineering. As a side note, this project was conceptualized after noticing the extensive number of deletions needed in metabolic engineering, and how often the engineering strategies required targeting paralogs from the same pathway, e.g. *PLB1*, *PLB2*, *PLB3*, *or FAA1*, *FAA4*, and I have since applied it for studies in human cell lines (189). There are other circumstances in which off-targeting might be desirable, e.g. finding synthetic lethality in cancer (190), or minimizing a genome (discussed in Chapter V) (191). Based on the current number of citations, our promiscuous gRNA tool has not been extensively utilized by other researchers. One reason might be the difficulty to use it and its low speed when targeting complex living organisms. Thus, I have decided to facilitate access to the information of existing promiscuous gRNA by developing an R Shiny application that contains a database of promiscuous gRNA from several organisms. This will remove the need to run our existing framework in R and hopefully allow more people to exploit promiscuous gRNA in their studies.

Chapter V: Conclusions & Perspectives



5.1 Future engineering strategies for the production of oleochemicals

Throughout the thesis, I have attempted to emphasize the technological diversity provided by CRISPR and apply it in the context of building efficient yeast cell factories for the production of oleochemicals, namely TAGs and FFAs in **Chapter II**, and 3-HP in **Chapter III**. In **Paper II**, we disrupted several metabolic fluxes, namely reactions involved in the fatty acid oxidation, the production of storage lipids, and the conversion of FFAs to fatty acyl-CoA, which, in the end, constrained the lipid metabolic network to its main components. The combined deletions minimized the complexity of the lipid metabolic network by disrupting many regulation mechanisms. The resulting strain, MLM1.0, efficiently redirected fatty acid fluxes toward phospholipids and yielded a strain accumulating high levels of both FFAs and phospholipids. While the FFA titers achieved in MLM1.0 remained below previous reports, it is worth to mention that the strain does not carry any additional overexpression of genes. It would confer an advantage when expressing heterologous pathways that further convert the FFAs to other valuable products, compared to the superior strains reported, as MLM1.0 could have more proteome mass available to allocate (119, 151, 152). For example, Leber and colleagues reported 2.2 $g \cdot L^{-1}$ of FFAs from an engineered strain designed to force fluxes towards TAG synthesis and hydrolysis, namely overexpression of *DGA1* and *TGL3*, in a *FAA1*, *FAA4*, *FAT1* and *FAA2*, *POX1*, and *PXA1* deletion strain (152). Our attempts to carry out similar strategy for MLM1.0, i.e. overexpressing endogenous phospholipases B, did not further increase the FFA levels. Notably, the expression of a thioesterase (‘TesA) from *E. coli* significantly improved FFA levels in MLM1.0 strain (**Paper V**: Fig. 5B). Thus, future strategies would need to elucidate ways to harness this accumulated pool of phospholipids, e.g. via the overexpression of efficient heterologous phospholipases, and fine-tuning of the flux from PA to PLs, as well as the acyl-CoA pool, e.g. via the overexpression of GPAT and LPAT genes (192). Notably, Yu and colleagues have recently achieved the highest FFA titers

reported by microbial fermentation (33.4g·L^{-1} by fed-batch fermentation) through extensive metabolic rewiring and adaptive laboratory evolution strategies (151). Sequential alteration of the metabolism was achieved by *(i)* the integration of heterologous pathways to produce FFAs, *(ii)* the deletion of the pyruvate decarboxylase encoding genes (*PDC1/5/6*) to disrupt the ethanol production from pyruvate, *(iii)* proper balancing of NADPH, ATP, and acetyl-CoA, and *(iv)* ultimately a directed laboratory evolution to improve cell growth on glucose. Because their strategy to achieve high titers extensively focus on pathways upstream acetyl-CoA, there is compatibility with our strategy developed in **Paper II**, where targeted genes were located downstream acetyl-CoA (Fig. V-1).

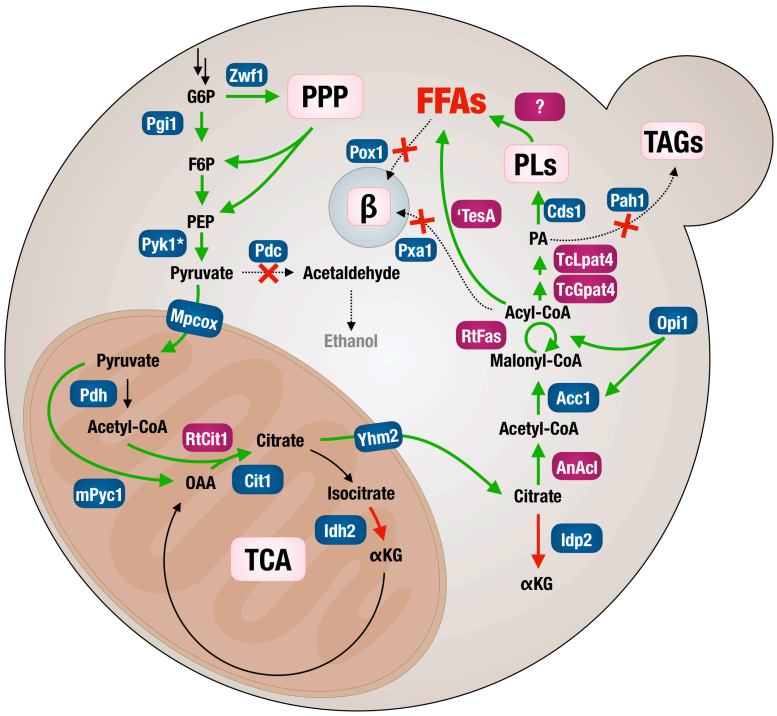


Figure V-1. Suggestions for future strategies for improving FFA production. Strategies inspired from the work presented in Paper II, combined with Yu et al. 2018 (151), and Wei et al. 2018 (192) reports. In blue: endogenous enzymes; purple: heterologous enzymes. Green arrow represents upregulated genes; red arrows: downregulated genes; crossed & dashed arrow: deleted genes; Pgi1: glucose-6-phosphate isomerase; Zwf1: glucose-6-phosphate 1-dehydrogenase; Pyk1*: mutated pyruvate kinase; Pdc: pyruvate decarboxylase; Mpcox: mitochondrial pyruvate carrier; RtCit1: citrate synthase from *Rhodosporidium toruloides*; Cit1: citrate synthase; mPyc1: mitochondrial pyruvate carboxylase 1; Idh2: mitochondrial isocitrate dehydrogenase 2; Yhm2: citrate/oxoglutarate carrier protein; Idp2: cytoplasmic isocitrate dehydrogenase; AnAcl: ATP citrate lyase from *Aspergillus nidulans*; Acc1: acetyl-CoA carboxylase 1; RtFas: fatty acid synthase from *R. toruloides*; *Opi1*: transcriptional repressor Opi1; 'TesA: acyl-CoA thioesterase A from *E. coli*; Cds1: phosphatidate cytidylyltransferase; Pah1: phosphatidic acid phosphohydrolase 1; Pox1: acyl-coenzyme A oxidase; Pxa1: peroxisomal long-chain fatty acid import protein 2; TcGpat4: glycerol-3-phosphate acyltransferase 4 from *Theobroma cacao*; LcPat4: lysophosphatidyl acyltransferase 4 from *Theobroma cacao*; ?: heterologous enzyme(s) harnessing PL pools, e.g. phospholipase B. TCA: tricarboxylic acid cycla; PLs: Phospholipids; PPP: pentose phosphate pathway; β: β-oxidation.

In **Paper II**, despite our efforts to engineer to *S. cerevisiae* to accumulate large amount of TAGs and ultimately achieving the highest titers reported so far, oleaginous yeasts, such as *Lipomyces starkeyi*, *R. toruloides*, *Y. lipolytica*, largely outperform *S. cerevisiae* (193). However, while we carried out many deletions, many of them turned out to have no effect on TAG levels (in a strain accumulating high levels of TAGs). Indeed, the main effects massively increasing TAGs levels were *(i)* the push-and-pull strategy, inspired from Runguphan and Keasling (121), where fatty-acid and TAG biosyntheses were upregulated by overexpressing *ACC1***, and *DGA1* + *PAH1*, respectively; and *(ii)* disruption of the main TAG hydrolysis genes, namely *TGL3*, *TGL4*, and *TGL4*, whose deletions were previously reported by Ploier and colleagues (194). Notably, among the different deletions carried out in the study, processing the acyl-CoA pool by deleting *PXA1*, which encodes a fatty acyl-CoA peroxisomal transporter, significantly increased the TAG pools. This result, together with the introduction of TesA in **Paper II** whose expression significantly increased FFA pools, might indicate a surplus of acyl-CoAs available for further processing. Future strategies to reach levels approximate to oleaginous yeasts would require extensive additional engineering, in addition to those mentioned for FFAs as they are compatible with TAGs production (Fig. V-1) (151, 192). For example, Teixeira and colleagues offered an approach to improve TAG levels based on modifying the structural mechanisms of the LD. In their study, they obtained a 138% increase from strain RF07 strain by overexpressing the human and fungal perilipins (PLIN3 and *MlPLP1*) which are involved in LD biogenesis, and deleting *ERD1* which encodes a predicted membrane protein required for lumenal ER protein retention (118).

5.2 The golden age of genetic engineering

Despite its only recent discovery, CRISPR/Cas9 technology has already been enhanced to the point of fulfilling most of the genome editing and gene regulation currently needed, ranging from the ability to perform multiple gene insertions, gene knockouts and the generation of combinatorial libraries to advanced fine-tuning of biosynthetic pathways. Thus, the industrial biotechnology field using metabolically engineered microbial cell factories has progressively been shifting from studies with few genetic modifications to highly complex engineered strains (188, 195–198). CRISPR has become an established tool for these microbial cell factories, and the field is observing more and more complex organisms being successfully engineered with this technology (199, 200). In **Chapter II**, CRISPR/Cas9 enabled us to rapidly construct strains, with our most engineered strain reaching eleven modifications (154). This technology, overall, sensibly increases the pace of the DBLT, as the building phase is empirically the longest phase (Fig. I-2). One can speculate that complete biosynthesis of complex compounds such as secondary metabolites will be more frequently reported using CRISPR technology to either integrate these complex pathways into a well-characterized organism or, directly genetically engineer the host organism. Decision on these two options, will, most likely, be based on the extent a model organism can be engineered to reach the desired function. The orthogonality of Cas

enzymes is relatively well established, and the main barrier for engineering these complex organisms remains the engineering efficiency. While the highly efficient HR machinery in *S. cerevisiae* has allowed CRISPR tools to be easily applied and developed, the advent of newly discovered *cas* genes with interesting properties could eventually circumvent the prominence of NHEJ in other organisms such as mammalian cells. For example, two independent labs recently reported a method for large DNA insertions using CRISPR-associated transposases, which would ultimately eliminate the dependency on the endogenous repair machinery to perform genetic engineering (201, 202). This is relevant in the context of producing oleochemicals in oleaginous yeast as many engineering challenges come from their predominant NHEJ repair machinery.

5.3 *Mundus Novus*

One aspect that remains underexploited in the metabolic engineering field is the utilization of CRISPR gRNA libraries for genome-wide high-throughput screening (203). In comparison to rational approaches, which predominate the field, appropriate screening platforms could enable unbiased and high-level perturbations of gene networks (204). In **Chapter III**, after characterizing dCas9-VPR as a robust tool for graded gene regulation (**Paper IV**), we harnessed the versatility of gRNA libraries with an encoded biosensor to identify optimal gene expression of (non-obvious) genes contributing to 3-HP production (**Paper V**). As with our study, model-guided strategies can provide a promising platform for inferring novel biological phenotypes, ultimately reducing the number of targeted genes. However, while our FBA simulations predicted candidate genes for screening and subsequently reduced the number of genes to target, most of the predicted genes were not represented among the enriched gRNAs. Here, potential issues arising from incorrect *in silico* predictions could come from the biomass composition set to be invariant at different growth rates, for which slight changes in the coefficients can have major effects on the obtained flux distribution. Even though the yeast GEM used in this study is extensively curated, many gaps might remain, e.g. dead-end metabolites and unconnected reactions (39, 205). Altogether, this highlights the need for a better scoring associated to these genes, and more importantly the relevance of the suggested genes.

Also, because of its unpredictable outcome, systematic screenings with combinatorial gRNA libraries remain challenging as gene interference is difficult to predict solely based on the gRNA position. In **Paper V**, to mitigate these disadvantages, we extensively covered each promoter targeted with up to 21 gRNAs. Notably, applying an extensive coverage could also help to elucidate functional regulatory sites in the promoter studied. Indeed, mapping the regulatory effect of each gRNA targeted to a transcription factor binding sites could eventually provide more insight into the importance of each transcription factors on the targeted promoter. Also, covering each promoter by up to 21 gRNAs allowed to, fortuitously, retrieve non-obvious transcriptional regulation that ultimately significantly increased 3-HP production, e.g. *SPO23* and *HTA1*. Likewise, Joung and colleagues targeted

promoters of more than 10,000 long non-coding RNAs using a CRISPR activation screening approach and identified several loci that drive anticancer drug resistance in melanoma cells (206). It would be of great interest to apply a similar strategy in yeast to systematically discover the functions of non-obvious elements that influence specific cellular processes. Conversely to the numerous large-scale studies performed in mammalian cells, the "yeast field" lacks a proper CRISPR database with sufficient knowledge needed to perform high-fidelity genetic engineering experiments, e.g. a list of all experimentally validated gRNAs used for gene knockout or gene regulation in yeast (207–210). The advantages brought by having such database would invaluably contribute to more complex combinatorial screenings, which would allow further elucidation of new emergent synergies (206).

Besides novel gene regulation, CRISPR tools can also be deployed for synthetic evolution of enzymes (211, 212). For example, combinatorial libraries generated by error-prone PCR can be efficiently integrated with CRISPR/Cas9, ultimately enabling the screening of hundreds of thousands of novel gene variants (213). By extension, base-editing systems such as those described in **Chapter IV**, would provide better control over which specific region to mutagenize compared to error-prone PCR (189, 214–216).

5.4 Radical redesign of the yeast genome

The Sc2.0 project, which ambitiously aim to synthesize and assemble the whole *S. cerevisiae* genome, has already been completed to more than one-third, with seven synthetic chromosomes completed (191). Sc2.0 chromosomes are designed to have all TAG stop codons replaced by TAA, as well as all repetitive elements removed with the exception of tRNAs, which are relocated to a new chromosome. Importantly, detailed physiologic assessments later revealed almost no phenotypic impact on the strains harboring these synthetic chromosomes, e.g. wild-type growth rates and same protein abundance (217). The complete synthesis has been estimated to be achieved before 2025 and will cost approximately US$1.25 million (191, 218). On the other side, using CRISPR/Cas9 technology, top-down approaches for minimizing the yeast genome have started to emerge (219, 220). Promiscuous gRNAs, presented in **Paper VII**, could be theoretically be employed for a same type of applications. Here, our algorithm predicted a single promiscuous gRNAs capable to target 148 of these elements at once, and with proper setup could be utilized for minimizing of redundant genetic material, as with the Sc2.0 project. Additionally, these promiscuous gRNAs could also enhance *in vivo* combinatorial chromosomal rearrangements to accelerate adaptive evolution of strains with improved genetic backgrounds (221).

5.5 Advancing CRISPR technologies through yeast engineering

In the first reports of Cas9 application in human cell lines, the authors cautioned on the specificity of achieving significant cleavage, as well as cautioned on the unwanted off-target effects often accompanied with genome engineering tool (47, 61). Scientists have since aspired to develop the next-generation of engineered Cas9 with higher fidelity, different PAM specificity, increased nuclease activity, as well as smaller protein length (57, 63, 65, 75, 222, 223). In this context, several studies have exploited yeast as a workhorse organism for *in vivo* directed evolution of protein engineering for CRISPR applications (184, 222). Casini and colleagues screened a mutagenized library of Cas9 for both on- and off-target activity using a yeast-based assay established with a promiscuous gRNA targeting *TRP1* to assess the on-target specificity, and *ADE2* for off-targeting (222). The resulting screening yielded four beneficial mutations whose combination results in an engineered Cas9 with drastically less off-target than the wild-type version while retaining sufficient on-target efficiency. Also, Abuddayeh and colleagues recently developed a unique C to U RNA editor, where the adenine deaminase domain of ADAR2 was mutagenized into a cytidine deaminase *via* multiple rounds of library screening in yeast (184). Indeed, in order to select increased C to U activity, they aimed to reverse mutations in the yeast *HIS3* marker. Deciphering genes and pathways impacting the specificity of the cleavage would also be of utmost interest. In this context, future strategies could employ yeast deletion or overexpression libraries as they provide a unique advantage compared to other tools.

Additionally, the advent of machine-learning algorithms for protein engineering could also enable the discovery of nontrivial Cas functions (224, 225). Notably, machine learning algorithms have already been employed to reliably predict NHEJ patterns from the repair of double-strand breaks in mammalian cells (225, 226). Finally, besides directed evolution, several studies have employed yeast to shed light on idiosyncratic patterns of HR repair mechanisms in heterozygous diploid cells (227, 228). Gorter de Vries and colleagues recently reported loss of heterozygosity of entire chromosome arms from unexpected HR events in heterozygous diploid yeast cells (228). Since these events could have dramatic consequences in human gene therapy or gene drive setups, preliminary yeast-based studies could eventually elucidate key factors to optimize these experiments.

Acknowledgements

First and foremost, I would like to thank my supervisor Jens Nielsen. My PhD studies have been a thrilling journey of scientific growth and personal development, and nothing would have been possible without your constant support and supervision throughout these years. Thank you for investing time and believing in me, and for being a thoughtful and compassionate mentor. I feel incredibly privileged to have worked with you.

To my two co-supervisors Florian and Verena for doing a great job at dealing with my impatience and for tolerating my incessant irruptions (especially at the beginning of my studies). Thanks for cheering me up when I was in doubt, I've learnt a lot from you.

Throughout my studies, I've had the chance to work with many great collaborators at SysBio. Thanks to Paulo for all the fruitful collaborations we've had. It has been a lot of fun partnering with you on all these projects. Gatto for all the scientific and non-scientific support. Thank you for continually destroying many of my naive ideas. My *frenduru* Ben for all the great discussions, trips, and sake. I'm glad that we managed to work on a project together. Julle for all the great discussions and fun we've had. Our attempt to start an automated strain platform was challenging but entertaining. Michi for your great assistance when I began in the lab. Iván for patiently teaching me about GEM. Christos, for all the excellent work on the Csy4 project and the CRISPR library. Angelo and Alex, I feel fortunate to have supervised you both at the same time. It has been gratifying. Dina for your many valuable feedbacks and support. Jon for your advice and assistance in the dry lab. Stefan for showing me the value of collaborative work.

I have been extremely privileged to have worked with amazing people outside SysBio. Thank you, Matthew Vander Heiden, for having me in your lab to work on cancer metabolism. I've had a fantastic time in Boston thanks to you and Brooke; I really felt welcomed. Special thanks to Ahmed and Keene for the all the great (late) work and discussion, and more importantly for introducing me to the finest American cuisine. Looking forward to seeing you both again this fall. Alisson, Peggy and Zhaoqi for allowing me to work on your amazing projects. I would like to thank George Church as well as members of his lab, Oscar and Cory, for our work on promiscuous gRNA applied in human cell lines. It is incredible how fast things have been moving for the past seven months, with each month bringing a new exciting challenge. I hope we will accomplish great things together. Finally, thanks to Michael Jensen and Emil Jensen from Jay Keasling's group. Our collaborative work on benchmarking CRISPR-dCas9 tools in yeast has been crucial for initiating several projects in this thesis.

SysBio is a fantastic international workplace, and I'd like to thank all the people who have made my stay entertaining and comfortable. Thanks to Ed, Tyler, Rui, Joakim, Christer, Martina, Sakda, Sylvain, Leif, Alexandra, Petri, Marie, Eugene, Andrea CL, Dimitra, Lucy, Olena, Ivan, Avlant, Yassi, Oliver, JC, Ponytail, De Fries, Carl, Johan, Siniša, Kate,

Rosemary, Promi, Anastasia, Martin, Aleksej, Shaq, Mihail, Andreas, Jing, Amir, Tao, Christoph, John, Rasool, Mia, Erica, Abder, Yongjin, Filip, Anne-Lise, Zhiwei, Stefan, Veronica, and Angelica.

Thanks to my friends Mat, Philip, Michpov, Ed, Jul, PA, Robbie, Jorgelindo, Fabian, Anna, Julia, and Viktor, from France and Sweden for all the entertaining moments along the way.

Last but not least, thanks to family and extended family. Most importantly, to my parents and sister for their eternal love and support. Ivana, thanks for sharing your life with me, and making mine awesome. I love you.

References

1. McGovern PE, et al. (2004) Fermented beverages of pre- and proto-historic China. *Proc Natl Acad Sci U S A* 101(51):17593–17598.

2. Cavalieri D, McGovern PE, Hartl DL, Mortimer R, Polsinelli M (2003) Evidence for S. cerevisiae Fermentation in Ancient Wine. *J Mol Evol* 57(1):S226–S232.

3. Pasteur L (1876) *Études sur la bière: ses maladies, causes qui les provoquent, procédé pour la rendre inaltérable; avec une théorie nouvelle de la fermentation* (Gauthier-Villars).

4. Sauer M (2016) Industrial production of acetone and butanol by fermentation—100 years later. *FEMS Microbiol Lett* 363(13). doi:10.1093/femsle/fnw134.

5. A ZW, A JZ, A HF, B BAP *Glycerol production by microbial fermentation: A review*.

6. Fleming A (1929) On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ. *Br J Exp Pathol* 10(3):226–236.

7. Freeman GG, Donald GMS (1957) Fermentation Processes Leading to Glycerol. *Appl Microbiol* 5(4):197–210.

8. Watson JD, Crick FHC (1953) Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* 171(4356):737.

9. Cohen SN, Chang ACY, Boyer HW, Helling RB (1973) Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proc Natl Acad Sci U S A* 70(11):3240–3244.

10. The pharmacological basis of therapeutics, fifth edition. Edited by Louis S. Goodman and Alfred Gilman. Macmillan, 866 Third Ave., New York, NY 10022, 1975. 1704 pp. 19 × 27 cm. Price $30.00 (1976) *J Pharm Sci* 65(5):781–781.

11. Gest H. (2004) The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society. *Notes Rec R Soc Lond* 58(2):187–201.

12. Koch RDM (1999) Koch l Methods for the study of pathogenic organisms.

13. Hansen EC (Emil C, 1842-1909 (1891) Recherches sur la physiologie et la morphology des ferments alcooliques. Available at: http://agris.fao.org/agris-search/search.do?recordID=US201300273671 [Accessed July 5, 2019].

14. Charles weizmann (1919) Available at: https://patents.google.com/patent/US1315585A/en [Accessed July 5, 2019].

15. Ereky Karl (1919) *Biotechnologie der Fleisch-, Fett-, und Milcherzeugung im landwirtschaftlichen Grossbetriebe: für naturwissenschaftlich gebildete Landwirte verfasst* (P. Parey, Berlin) Available at: https://catalog.hathitrust.org/Record/006798043 [Accessed July 5, 2019].

16. Lombardino JG, Erhart C (2000) A BRIEF HISTORY OF PFIZER CENTRAL RESEARCH.

17. Gould K (2016) Antibiotics: from prehistory to the present day. *J Antimicrob Chemother* 71(3):572–575.

18. Beadle GW, Tatum EL (1941) Genetic Control of Biochemical Reactions in Neurospora. *Proc Natl Acad Sci U S A* 27(11):499–506.

19. Goeddel DV, et al. (1979) Expression in Escherichia coli of chemically synthesized genes for human insulin. *Proc Natl Acad Sci* 76(1):106–110.

20. Metabolic Engineering: Principles and Methodologies - George Stephanopoulos, Aristos A. Aristidou, Jens Nielsen - Google Books Available at: https://books.google.se/books?hl=en&lr=&id=9mGzkso4NVQC&oi=fnd&pg=PP1&dq=info:2yQFm1ZF7UEJ:scholar.google.com&ots=sl1d8DQ3v8&sig=TeGJtzJLF-AoJQgGfP9vT5JOAGg&redir_esc=y#v=onepage&q&f=false [Accessed June 6, 2019].

21. Goffeau A, et al. (1996) Life with 6000 Genes. *Science* 274(5287):546–567.

22. Ro D-K, et al. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440(7086):940.

23. Tong AHY, et al. (2001) Systematic Genetic Analysis with Ordered Arrays of Yeast Deletion Mutants. *Science* 294(5550):2364–2368.

24. Jinek M, et al. (2012) A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337(6096):816–821.

25. Annaluru N, et al. (2014) Total Synthesis of a Functional Designer Eukaryotic Chromosome. *Science* 344(6179):55–58.

26. Galanie S, Thodey K, Trenchard IJ, Interrante MF, Smolke CD (2015) Complete biosynthesis of opioids in yeast. *Science* 349(6252):1095–1100.

27. Luo X, et al. (2019) Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature*. doi:10.1038/s41586-019-0978-9.

28. Hillson N, et al. (2019) Building a global alliance of biofoundries. *Nat Commun* 10(1):2040.

29. Pauling L, Corey RB, Branson HR (1951) The Structure of Proteins. *Proc Natl Acad Sci U S A* 37(4):205–211.

30. Muller HJ (1927) ARTIFICIAL TRANSMUTATION OF THE GENE. *Science* 66(1699):84–87.

31. Barnett JA (2007) A history of research on yeasts 10: foundations of yeast genetics1. *Yeast* 24(10):799–845.

32. Winge Ö (1952) The genetic situation concerning fermentation in yeasts. *Heredity* 6(2):263–269.

33. Winge O, Roberts C (1950) The polymeric genes for maltose fermentation in yeast and their mutability. *Comptes Rendus Trav Lab Carlsberg Ser Physiol* 25:35–83.

34. Russo E (2003) Special Report: The birth of biotechnology. *Nature* 421(6921):456.

35. Thim L, Hansen MT, Sørensen AR (1987) Secretion of human insulin by a transformed yeast cell. *FEBS Lett* 212(2):307–312.

36. Bailey JE (1991) Toward a science of metabolic engineering. *Science* 252(5013):1668–1675.

37. Campbell K, Xia J, Nielsen J (2017) The Impact of Systems Biology on Bioprocessing. *Trends Biotechnol* 35(12):1156–1168.

38. Zhang C, Hua Q (2016) Applications of Genome-Scale Metabolic Models in Biotechnology and Systems Medicine. *Front Physiol* 6. doi:10.3389/fphys.2015.00413.

39. Sánchez BJ, Nielsen J (2015) Genome scale models of yeast: towards standardized evaluation and consistent omic integration. *Integr Biol* 7(8):846–858.

40. Oberhardt MA, Palsson BØ, Papin JA (2009) Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* 5(1):320.

41. Alberts B, et al. (2002) *Molecular Biology of the Cell* (Garland Science). 4th Ed.

42. Gibson DG, et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5):343–345.

43. Cheng F, Tang X-L, Kardashliev T (2018) Transcription Factor-Based Biosensors in High-Throughput Screening: Advances and Applications. *Biotechnol J* 13(7):1700648.

44. Petzold CJ, Chan LJG, Nhan M, Adams PD (2015) Analytics for metabolic engineering. *Front Bioeng Biotechnol* 3. doi:10.3389/fbioe.2015.00135.

45. Siewers V (2014) An Overview on Selection Marker Genes for Transformation of Saccharomyces cerevisiae. *Yeast Metabolic Engineering: Methods and Protocols*, Methods in Molecular Biology., ed Mapelli V (Springer New York, New York, NY), pp 3–15.

46. Jinek M, et al. (2013) RNA-programmed genome editing in human cells. *eLife* 2. doi:10.7554/eLife.00471.

47. Mali P, et al. (2013) RNA-Guided Human Genome Engineering via Cas9. *Science* 339(6121):823–826.

48. Ferreira R, David F, Nielsen J (2018) Advancing biotechnology with CRISPR/Cas9: recent applications and patent landscape. *J Ind Microbiol Biotechnol* 45(7):467–480.

49. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol* 169(12):5429–5433.

50. Mojica FJM, Rodriguez-Valera F (2016) The discovery of CRISPR in archaea and bacteria. *FEBS J* 283(17):3162–3169.

51. Ishino Y, Krupovic M, Forterre P (2018) History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. *J Bacteriol* 200(7). doi:10.1128/JB.00580-17.

52. Jackson SA, et al. (2017) CRISPR-Cas: Adapting to change. *Science* 356(6333):eaal5056.

53. S. Makarova K, et al. (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9(6):467–477.

54. Makarova KS, Wolf YI, Koonin EV (2018) Classification and Nomenclature of CRISPR-Cas Systems: Where from Here? *CRISPR J* 1(5):325–336.

55. Doudna JA, Charpentier E (2014) The new frontier of genome engineering with CRISPR-Cas9. *Science* 346(6213):1258096.

56. Hsu PD, Lander ES, Zhang F (2014) Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* 157(6):1262–1278.

57. Nishimasu H, et al. (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361(6408):1259–1262.

58. Nishimasu H, et al. (2014) Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* 156(5):935–949.

59. Jiang F, Doudna JA (2017) CRISPR–Cas9 Structures and Mechanisms. *Annu Rev Biophys* 46(1):505–529.

60. Shah SA, Erdmann S, Mojica FJM, Garrett RA (2013) Protospacer recognition motifs. *RNA Biol* 10(5):891–899.

61. Cong L, et al. (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 339(6121):819–823.

62. Wang H, La Russa M, Qi LS (2016) CRISPR/Cas9 in Genome Editing and Beyond. *Annu Rev Biochem* 85(1):227–264.

63. Hu JH, et al. (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556(7699):57–63.

64. Slaymaker IM, et al. (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science* 351(6268):84–88.

65. Pickar-Oliver A, Gersbach CA (2019) The next generation of CRISPR–Cas technologies and applications. *Nat Rev Mol Cell Biol*:1.

66. Qi LS, et al. (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* 152(5):1173–1183.

67. Perez-Pinera P, et al. (2013) RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nat Methods* 10(10):973–976.

68. Farzadfard F, Perli SD, Lu TK (2013) Tunable and Multifunctional Eukaryotic Transcription Factors Based on CRISPR/Cas. *ACS Synth Biol* 2(10):604–613.

69. Deaner M, Alper HS (2017) Systematic testing of enzyme perturbation sensitivities via graded dCas9 modulation in Saccharomyces cerevisiae. *Metab Eng* 40:14–22.

70. Jensen ED, et al. (2017) Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. *Microb Cell Factories* 16(1):46.

71. Chavez A, et al. (2015) Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* 12(4):326–328.

72. Maeder ML, et al. (2013) CRISPR RNA–guided activation of endogenous human genes. *Nat Methods* 10(10):977–979.

73. Zetsche B, et al. (2015) Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* 163(3):759–771.

74. Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E (2016) The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 532(7600):517–521.

75. Liu J-J, et al. (2019) CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* 566(7743):218.

76. Wong N, Liu W, Wang X (2015) WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. *Genome Biol* 16(1):218.

77. Tsai SQ, et al. (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33(2):187–197.

78. Horlbeck MA, et al. (2016) Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *eLife* 5. doi:10.7554/eLife.12677.

79. Nielsen J (2013) Production of biopharmaceutical proteins by yeast. *Bioengineered* 4(4):207–211.

80. Nielsen J (2019) Yeast Systems Biology: Model Organism and Cell Factory. *Biotechnol J*:1800421.

81. Nielsen J (2017) Systems Biology of Metabolism. *Annu Rev Biochem* 86(1):245–275.

82. Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction (1992) *J Cell Biol* 119(2):301–311.

83. Petranovic D, Tyo K, Vemuri GN, Nielsen J (2010) Prospects of yeast systems biology for human health: integrating lipid, protein and energy metabolism: Yeast systems biology for human health. *FEMS Yeast Res* 10(8):1046–1059.

84. Scherens B, Goffeau A (2004) The uses of genome-wide yeast mutant collections. *Genome Biol* 5(7):229.

85. Srivas R, et al. (2016) A network of conserved synthetic lethal interactions for exploration of precision cancer therapy. *Mol Cell* 63(3):514–525.

86. Ferreira R, Limeta A, Nielsen J (2019) Tackling Cancer with Yeast-Based Technologies. *Trends Biotechnol* 37(6):592–603.

87. France hits all-time record temperature of 45.9C (2019) Available at: https://www.bbc.com/news/world-europe-48795264 [Accessed July 1, 2019].

88. AR5 Climate Change 2013: The Physical Science Basis — IPCC Available at: https://www.ipcc.ch/report/ar5/wg1/ [Accessed July 1, 2019].

89. Power SB, Delage FPD (2019) Setting and smashing extreme temperature records over the coming century. *Nat Clim Change* 9(7):529.

90. Media Release: Nature's Dangerous Decline 'Unprecedented'; Species Extinction Rates 'Accelerating' | IPBES Available at: https://www.ipbes.net/news/Media-Release-Global-Assessment [Accessed July 1, 2019].

91. Gynther I, Waller N Confirmation of the extinction of the Bramble Cay melomys. 65.

92. Urban MC (2015) Accelerating extinction risk from climate change. *Science* 348(6234):571–573.

93. Johnston ER, et al. (2019) Responses of tundra soil microbial communities to half a decade of experimental warming at two critical depths. *Proc Natl Acad Sci*:201901307.

94. Turetsky MR, et al. (2019) Permafrost collapse is accelerating carbon release. *Nature* 569(7754):32.

95. Liao JC, Mi L, Pontrelli S, Luo S (2016) Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nat Rev Microbiol* 14(5):288–304.

96. Oils and Fats as Renewable Raw Materials in Chemistry - Biermann - 2011 - Angewandte Chemie International Edition - Wiley Online Library Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/anie.201002767 [Accessed July 4, 2019].

97. Murphy DJ (2016) Plant Storage Lipids. *ELS* (American Cancer Society), pp 1–7.

98. Yan Q, Pfleger BF (2019) Revisiting metabolic engineering strategies for microbial synthesis of oleochemicals. *Metab Eng*:S1096717619301132.

99. Rabiu A, Elias S, Oyekola O (2018) Oleochemicals from Palm Oil for the Petroleum Industry. *Palm Oil*. doi:10.5772/intechopen.76771.

100. Vijay V, Pimm SL, Jenkins CN, Smith SJ (2016) The Impacts of Oil Palm on Recent Deforestation and Biodiversity Loss. *PLOS ONE* 11(7):e0159668.

101. Marella ER, Holkenbrink C, Siewers V, Borodina I (2018) Engineering microbial fatty acid metabolism for biofuels and biochemicals. *Curr Opin Biotechnol* 50:39–46.

102. Pfleger BF, Gossing M, Nielsen J (2015) Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metab Eng* 29:1–11.

103. Schüller HJ, Hahn A, Tröster F, Schütz A, Schweizer E (1992) Coordinate genetic control of yeast fatty acid synthase genes FAS1 and FAS2 by an upstream activation site common to genes involved in membrane lipid biosynthesis. *EMBO J* 11(1):107–114.

104. Hasslacher M, Ivessa AS, Paltauf F, Kohlwein SD (1993) Acetyl-CoA carboxylase from yeast is an essential enzyme and is regulated by factors that control phospholipid metabolism. *J Biol Chem* 268(15):10946–10952.

105. Loewen CJR, et al. (2004) Phospholipid Metabolism Regulated by a Transcription Factor Sensing Phosphatidic Acid. *Science* 304(5677):1644–1647.

106. Zhang M, Galdieri L, Vancura A (2013) The Yeast AMPK Homolog SNF1 Regulates Acetyl Coenzyme A Homeostasis and Histone Acetylation. *Mol Cell Biol* 33(23):4701–4717.

107. Tehlivets O, Scheuringer K, Kohlwein SD (2007) Fatty acid synthesis and elongation in yeast. *Biochim Biophys Acta* 1771(3):255–270.

108. Stukey JE, McDonough VM, Martin CE (1990) The OLE1 gene of Saccharomyces cerevisiae encodes the delta 9 fatty acid desaturase and can be functionally replaced by the rat stearoyl-CoA desaturase gene. *J Biol Chem* 265(33):20144–20149.

109. Oh C-S, Toke DA, Mandala S, Martin CE (1997) ELO2 and ELO3, Homologues of theSaccharomyces cerevisiae ELO1 Gene, Function in Fatty Acid Elongation and Are Required for Sphingolipid Formation. *J Biol Chem* 272(28):17376–17384.

110. Toke DA, Martin CE (1996) Isolation and Characterization of a Gene Affecting Fatty Acid Elongation in Saccharomyces cerevisiae. *J Biol Chem* 271(31):18413–18422.

111. Yu T, et al. (2017) Metabolic engineering of *Saccharomyces cerevisiae* for production of very long chain fatty acid-derived chemicals. *Nat Commun* 8:15587.

112. Zhu Z, et al. (2017) Expanding the product portfolio of fungal type I fatty acid synthases. *Nat Chem Biol* 13(4):360–362.

113. Bergenholm D, Gossing M, Wei Y, Siewers V, Nielsen J (2018) Modulation of saturation and chain length of fatty acids in Saccharomyces cerevisiae for production of cocoa butter-like lipids. *Biotechnol Bioeng* 115(4):932–942.

114. Black PN, DiRusso CC (2007) Yeast acyl-CoA synthetases at the crossroads of fatty acid metabolism and regulation. *Biochim Biophys Acta BBA - Mol Cell Biol Lipids* 1771(3):286–298.

115. Daum G (2004) Introduction: Lipids: cellular glue...or are they more than that? *Lipid Metabolism and Membrane Biogenesis*, ed Daum G (Springer Berlin Heidelberg, Berlin, Heidelberg), pp 1–3.

116. Yang H, et al. (1996) Sterol Esterification in Yeast: A Two-Gene Process. *Science* 272(5266):1353–1356.

117. Athenstaedt K, Daum G (2006) The life cycle of neutral lipids: synthesis, storage and degradation. *Cell Mol Life Sci CMLS* 63(12):1355–1369.

118. Teixeira PG, Ferreira R, Zhou YJ, Siewers V, Nielsen J (2017) Dynamic regulation of fatty acid pools for improved production of fatty alcohols in Saccharomyces cerevisiae. *Microb Cell Factories* 16(1):45.

119. Zhou YJ, et al. (2016) Production of fatty acid-derived oleochemicals and biofuels by synthetic yeast cell factories. *Nat Commun* 7:11709.

120. Buijs NA, Zhou YJ, Siewers V, Nielsen J (2015) Long-chain alkane production by the yeast Saccharomyces cerevisiae. *Biotechnol Bioeng* 112(6):1275–1279.

121. Runguphan W, Keasling JD (2014) Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals. *Metab Eng* 21:103–113.

122. Shi S, Valle-Rodríguez JO, Khoomrung S, Siewers V, Nielsen J (2012) Functional expression and characterization of five wax ester synthases in Saccharomyces cerevisiae and their utility for biodiesel production. *Biotechnol Biofuels* 5:7.

123. Stöveken T, Kalscheuer R, Malkus U, Reichelt R, Steinbüchel A (2005) The Wax Ester Synthase/Acyl Coenzyme A:Diacylglycerol Acyltransferase from Acinetobacter sp. Strain ADP1: Characterization of a Novel Type of Acyltransferase. *J Bacteriol* 187(4):1369–1376.

124. Alvarez HM (2010) Biotechnological Production and Significance of Triacylglycerols and Wax Esters. *Handbook of Hydrocarbon and Lipid Microbiology*, ed Timmis KN (Springer Berlin Heidelberg, Berlin, Heidelberg), pp 2995–3002.

125. Wenning L, Yu T, David F, Nielsen J, Siewers V (2017) Establishing very long-chain fatty alcohol and wax ester biosynthesis in Saccharomyces cerevisiae. *Biotechnol Bioeng* 114(5):1025–1035.

126. Wenning L, et al. (2019) Increasing jojoba-like wax ester production in Saccharomyces cerevisiae by enhancing very long-chain, monounsaturated fatty acid synthesis. *Microb Cell Factories* 18(1):49.

127. Sarria S, Kruyer NS, Peralta-Yahya P (2017) Microbial synthesis of medium-chain chemicals from renewables. *Nat Biotechnol* 35(12):1158–1166.

128. Choi YJ, Lee SY (2013) Microbial production of short-chain alkanes. *Nature* 502(7472):571–574.

129. Ejsing CS, et al. (2009) Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry. *Proc Natl Acad Sci* 106(7):2136–2141.

130. Zhu Z, et al. (2017) Enabling the synthesis of medium chain alkanes and 1-alkenes in yeast. *Metab Eng* 44:81–88.

131. Zhang P, Chen Q, Fu G, Xia linglin, Hu X (2019) Regulation and metabolic engineering strategies for permeases of Saccharomyces cerevisiae. *World J Microbiol Biotechnol* 35(7):112.

132. Krivoruchko A, Nielsen J (2015) Production of natural products through metabolic engineering of Saccharomyces cerevisiae. *Curr Opin Biotechnol* 35:7–15.

133. Krivoruchko A, Zhang Y, Siewers V, Chen Y, Nielsen J (2015) Microbial acetyl-CoA metabolism and metabolic engineering. *Metab Eng* 28:28–42.

134. Hu Y, Zhu Z, Nielsen J, Siewers V (2018) Heterologous transporter expression for improved fatty alcohol secretion in yeast. *Metab Eng* 45:51–58.

135. Hu Y, Zhu Z, Nielsen J, Siewers V (2019) Engineering *Saccharomyces cerevisiae* cells for production of fatty acid-derived biofuels and chemicals. *Open Biol* 9(5):190049.

136. Hammer SK, Avalos JL (2017) Harnessing yeast organelles for metabolic engineering. *Nat Chem Biol* 13(8):823–832.

137. Guo Z, Zhang L, Ding Z, Shi G (2011) Minimization of glycerol synthesis in industrial ethanol yeast without influencing its fermentation performance. *Metab Eng* 13(1):49–59.

138. de Jong BW, Shi S, Siewers V, Nielsen J (2014) Improved production of fatty acid ethyl esters in Saccharomyces cerevisiae through up-regulation of the ethanol degradation pathway and expression of the heterologous phosphoketolase pathway. *Microb Cell Factories* 13(1):39.

139. Athenstaedt K, Daum G (2003) YMR313c/TGL3 Encodes a Novel Triacylglycerol Lipase Located in Lipid Particles of Saccharomyces cerevisiae. *J Biol Chem* 278(26):23317–23323.

140. Klein I, et al. (2016) Regulation of the yeast triacylglycerol lipases Tgl4p and Tgl5p by the presence/absence of nonpolar lipids. *Mol Biol Cell* 27(13):2014–2024.

141. Köffel R, Tiwari R, Falquet L, Schneiter R (2005) The Saccharomyces cerevisiae YLL012/YEH1, YLR020/YEH2, and TGL1 Genes Encode a Novel Family of Membrane-Anchored Lipases That Are Required for Steryl Ester Hydrolysis. *Mol Cell Biol* 25(5):1655–1668.

142. Lee KS, et al. (1994) The Saccharomyces cerevisiae PLB1 gene encodes a protein required for lysophospholipase and phospholipase B activity. *J Biol Chem* 269(31):19725–19730.

143. Merkel O, et al. (1999) Characterization and Function in Vivo of Two Novel Phospholipases B/Lysophospholipases fromSaccharomyces cerevisiae. *J Biol Chem* 274(40):28121–28127.

144. Ferreira R, et al. (2018) Metabolic engineering of Saccharomyces cerevisiae for overproduction of triacylglycerols. *Metab Eng Commun* 6:22–27.

145. Kohlwein SD, Veenhuis M, Klei IJ van der (2013) Lipid Droplets and Peroxisomes: Key Players in Cellular Lipid Homeostasis or A Matter of Fat—Store 'em Up or Burn 'em Down. *Genetics* 193(1):1–50.

146. Czabany T, Athenstaedt K, Daum G (2007) Synthesis, storage and degradation of neutral lipids in yeast. *Biochim Biophys Acta BBA - Mol Cell Biol Lipids* 1771(3):299–309.

147. Friedlander J, et al. (2016) Engineering of a high lipid producing Yarrowia lipolytica strain. *Biotechnol Biofuels* 9(1):77.

148. Dulermo T, Nicaud J-M (2011) Involvement of the G3P shuttle and β-oxidation pathway in the control of TAG synthesis and lipid accumulation in Yarrowia lipolytica. *Metab Eng* 13(5):482–491.

149. Carman GM, Henry SA (2007) Phosphatidic Acid Plays a Central Role in the Transcriptional Regulation of Glycerophospholipid Synthesis in Saccharomyces cerevisiae. *J Biol Chem* 282(52):37293–37297.

150. Fakas S, et al. (2011) Phosphatidate Phosphatase Activity Plays Key Role in Protection against Fatty Acid-induced Toxicity in Yeast. *J Biol Chem* 286(33):29074–29085.

151. Yu T, et al. (2018) Reprogramming Yeast Metabolism from Alcoholic Fermentation to Lipogenesis. *Cell* 174(6):1549-1558.e14.

152. Leber C, Polson B, Fernandez-Moya R, Da Silva NA (2015) Overproduction and secretion of free fatty acids through disrupted neutral lipid recycle in Saccharomyces cerevisiae. *Metab Eng* 28:54–62.

153. van Hoogevest P, Wendel A (2014) The use of natural and synthetic phospholipids as pharmaceutical excipients. *Eur J Lipid Sci Technol* 116(9):1088–1107.

154. Ferreira R, Teixeira PG, Siewers V, Nielsen J (2018) Redirection of lipid flux toward phospholipids in yeast increases fatty acid turnover and secretion. *Proc Natl Acad Sci* 115(6):1262–1267.

155. Valle-Rodríguez JO, Shi S, Siewers V, Nielsen J (2014) Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid ethyl esters, an advanced biofuel, by eliminating non-essential fatty acid utilization pathways. *Appl Energy* 115:226–232.

156. Zalatan JG, et al. (2015) Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell* 160(1):339–350.

157. Smith JD, et al. (2016) Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. *Genome Biol* 17(1):45.

158. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res* 42(W1):W401–W407.

159. Qiao K, et al. (2015) Engineering lipid overproduction in the oleaginous yeast Yarrowia lipolytica. *Metab Eng* 29:56–65.

160. Gossing M, Smialowska A, Nielsen J (2018) Impact of forced fatty acid synthesis on metabolism and physiology of Saccharomyces cerevisiae. *FEMS Yeast Res* 18(8). doi:10.1093/femsyr/foy096.

161. Braun CJ, et al. (2016) Versatile in vivo regulation of tumor phenotypes by dCas9-mediated transcriptional perturbation. *Proc Natl Acad Sci* 113(27):E3892–E3900.

162. Williams TC, Pretorius IS, Paulsen IT (2016) Synthetic Evolution of Metabolic Productivity Using Biosensors. *Trends Biotechnol* 34(5):371–381.

163. David F, Nielsen J, Siewers V (2016) Flux Control at the Malonyl-CoA Node through Hierarchical Dynamic Pathway Regulation in *Saccharomyces cerevisiae*. *ACS Synth Biol* 5(3):224–233.

164. Johnson AO, et al. (2017) Design and application of genetically-encoded malonyl-CoA biosensors for metabolic engineering of microbial cell factories. *Metab Eng* 44:253–264.

165. Aung HW, Henry SA, Walker LP (2013) Revising the Representation of Fatty Acid, Glycerolipid, and Glycerophospholipid Metabolism in the Consensus Model of Yeast Metabolism. *Ind Biotechnol* 9(4):215–228.

166. Edwards JS, Palsson BO (1999) Systems Properties of the Haemophilus influenzaeRd Metabolic Genotype. *J Biol Chem* 274(25):17410–17416.

167. Orth JD, Thiele I, Palsson BØ (2010) What is flux balance analysis? *Nat Biotechnol* 28(3):245–248.

168. Shi S, Chen Y, Siewers V, Nielsen J (2014) Improving Production of Malonyl Coenzyme A-Derived Metabolites by Abolishing Snf1-Dependent Regulation of Acc1. *mBio* 5(3). doi:10.1128/mBio.01130-14.

169. Jayachandran C, Athiyaman BP, Sankaranarayanan M (2017) Cofactor engineering improved CALB production in Pichia pastoris through heterologous expression of NADH oxidase and adenylate kinase. *PLOS ONE* 12(7):e0181370.

170. Mayer FV, et al. (2011) ADP Regulates SNF1, the Saccharomyces cerevisiae Homolog of AMP-Activated Protein Kinase. *Cell Metab* 14(5):707–714.

171. Sakai Y, Rogi T, Yonehara T, Kato N, Tani Y (1994) High–Level ATP Production by a Genetically–Engineered Candida Yeast. *Nat Biotechnol* 12(3):291–293.

172. Aranda A, del Olmo Marcel l (2003) Response to acetaldehyde stress in the yeastSaccharomyces cerevisiae involves a strain-dependent regulation of severalALD genes and is mediatedby the general stress response pathway. *Yeast* 20(8):747–759.

173. Tevzadze GG, Pierce JV, Esposito RE (2007) Genetic Evidence for a SPO1-Dependent Signaling Pathway Controlling Meiotic Progression in Yeast. *Genetics* 175(3):1213–1227.

174. Lawson JE, Behal RH, Reed LJ (1991) Disruption and mutagenesis of the Saccharomyces cerevisiae PDX1 gene encoding the protein X component of the pyruvate dehydrogenase complex. *Biochemistry* 30(11):2834–2839.

175. Jakočiūnas T, et al. (2015) Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. *Metab Eng* 28:213–222.

176. Bao Z, et al. (2018) Genome-scale engineering of *Saccharomyces cerevisiae* with single-nucleotide precision. *Nat Biotechnol* 36(6):505–508.

177. Strohkendl I, Saifuddin FA, Rybarski JR, Finkelstein IJ, Russell R (2018) Kinetic Basis for DNA Target Specificity of CRISPR-Cas12a. *Mol Cell* 71(5):816-824.e3.

178. Świat MA, et al. (2017) FnCpf1: a novel and efficient genome editing tool for Saccharomyces cerevisiae. *Nucleic Acids Res* 45(21):12585–12598.

179. Liao C, et al. (2019) Modular one-pot assembly of CRISPR arrays enables library generation and reveals factors influencing crRNA biogenesis. *Nat Commun* 10(1):2948.

180. Smargon AA, et al. (2017) Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol Cell* 65(4):618-630.e7.

181. Moon SB, Kim DY, Ko J-H, Kim J-S, Kim Y-S (2019) Improving CRISPR Genome Editing by Engineering Guide RNAs. *Trends Biotechnol*. doi:10.1016/j.tibtech.2019.01.009.

182. Abudayyeh OO, et al. (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353(6299):aaf5573.

183. Cox DBT, et al. (2017) RNA editing with CRISPR-Cas13. *Science* 358(6366):1019–1027.

184. Abudayyeh OO, et al. (2019) A cytosine deaminase for programmable single-base RNA editing. *Science*:eaax7063.

185. Jing X, et al. (2018) Implementation of the CRISPR-Cas13a system in fission yeast and its repurposing for precise RNA editing. *Nucleic Acids Res* 46(15):e90–e90.

186. Mohanraju P, et al. (2016) Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353(6299):aad5147.

187. Ferry QRV, Lyutova R, Fulga TA (2017) Rational design of inducible CRISPR guide RNAs for de novo assembly of transcriptional programs. *Nat Commun* 8. doi:10.1038/ncomms14633.

188. McCarty NS, Shaw WM, Ellis T, Ledesma-Amaro R (2019) Rapid Assembly of gRNA Arrays *via* Modular Cloning in Yeast. *ACS Synth Biol* 8(4):906–910.

189. Smith CJ, et al. (2019) Enabling large-scale genome editing by reducing DNA nicking. *bioRxiv*. doi:10.1101/574020.

190. O'Neil NJ, Bailey ML, Hieter P (2017) Synthetic lethality and cancer. *Nat Rev Genet* 18(10):613–623.

191. Richardson SM, et al. (2017) Design of a synthetic yeast genome. *Science* 355(6329):1040–1044.

192. Wei Y, Bergenholm D, Gossing M, Siewers V, Nielsen J (2018) Expression of cocoa genes in Saccharomyces cerevisiae improves cocoa butter production. *Microb Cell Factories* 17(1):11.

193. Spagnuolo M, Yaguchi A, Blenner M (2019) Oleaginous yeast for biofuel and oleochemical production. *Curr Opin Biotechnol* 57:73–81.

194. Ploier B, et al. (2013) Screening for Hydrolytic Enzymes Reveals Ayr1p as a Novel Triacylglycerol Lipase in Saccharomyces cerevisiae. *J Biol Chem* 288(50):36061–36072.

195. Gander MW, Vrana JD, Voje WE, Carothers JM, Klavins E (2017) Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. *Nat Commun* 8:15459.

196. Roy KR, et al. (2018) Multiplexed precision genome editing with trackable genomic barcodes in yeast. *Nat Biotechnol* 36(6):512–520.

197. Guo X, et al. (2018) High-throughput creation and functional profiling of DNA sequence variant libraries using CRISPR–Cas9 in yeast. *Nat Biotechnol* 36(6):540–546.

198. Shi S, Liang Y, Zhang MM, Ang EL, Zhao H (2016) A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in Saccharomyces cerevisiae. *Metab Eng* 33:19–27.

199. Cai P, Gao J, Zhou Y (2019) CRISPR-mediated genome editing in non-conventional yeasts for biotechnological applications. *Microb Cell Factories* 18(1):63.

200. Martin-Laffon J, Kuntz M, Ricroch AE (2019) Worldwide CRISPR patent landscape shows strong geographical biases. *Nat Biotechnol* 37(6):613.

201. Strecker J, et al. (2019) RNA-guided DNA insertion with CRISPR-associated transposases. *Science*:eaax9181.

202. Klompe SE, Vo PLH, Halpin-Healy TS, Sternberg SH (2019) Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration. *Nature*:1.

203. Lian J, Mishra S, Zhao H (2018) Recent advances in metabolic engineering of Saccharomyces cerevisiae: New tools and their applications. *Metab Eng* 50:85–108.

204. Chen Y-C, et al. (2017) Randomized CRISPR-Cas transcriptional perturbation screening reveals protective genes against alpha-synuclein toxicity. *Mol Cell* 68(1):247-257.e5.

205. Kerkhoven EJ, Lahtvee P-J, Nielsen J (2015) Applications of computational modeling in metabolic engineering of yeast. *FEMS Yeast Res* 15(1):1–13.

206. Joung J, et al. (2017) Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* 548(7667):343–346.

207. Doench JG, et al. (2014) Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nat Biotechnol* 32(12):1262–1267.

208. Sanson KR, et al. (2018) Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat Commun* 9(1):5416.

209. Hart T, et al. (2017) Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3 Genes Genomes Genet* 7(8):2719–2727.

210. Haeussler M, et al. (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* 17(1):148.

211. Simon AJ, d'Oelsnitz S, Ellington AD (2019) Synthetic evolution. *Nat Biotechnol*. doi:10.1038/s41587-019-0157-4.

212. Halperin SO, et al. (2018) CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature* 560(7717):248–252.

213. Jakočiūnas T, Pedersen LE, Lis AV, Jensen MK, Keasling JD (2018) CasPER, a method for directed evolution in genomic contexts using mutagenesis and CRISPR/Cas9. *Metab Eng* 48:288–296.

214. Ma Y, et al. (2016) Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat Methods* 13(12):1029–1035.

215. Gaudelli NM, et al. (2017) Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 551(7681):464–471.

216. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533(7603):420–424.

217. van der Sloot A, Tyers M (2017) Synthetic Genomics: Rewriting the Genome Chromosome by Chromosome. *Mol Cell* 66(4):441–443.

218. Ellis T What is synthetic genomics anyway? *Synth Biol*:4.

219. Shao Y, et al. (2018) Creating a functional single-chromosome yeast. *Nature* 560(7718):331.

220. Luo J, Sun X, Cormack BP, Boeke JD (2018) Karyotype engineering by chromosome fusion leads to reproductive isolation in yeast. *Nature* 560(7718):392.

221. Blount BA, et al. (2018) Rapid host strain improvement by in vivo rearrangement of a synthetic yeast chromosome. *Nat Commun* 9(1):1932.

222. Casini A, et al. (2018) A highly specific SpCas9 variant is identified by *in vivo* screening in yeast. *Nat Biotechnol* 36(3):265–271.

223. Vakulskas CA, et al. (2018) A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat Med* 24(8):1216.

224. Yang KK, Wu Z, Arnold FH (2019) Machine-learning-guided directed evolution for protein engineering. *Nat Methods*. doi:10.1038/s41592-019-0496-6.

225. Shen MW, et al. (2018) Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* 563(7733):646.

226. Allen F, et al. (2019) Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat Biotechnol* 37(1):64–72.

227. DiCarlo JE, Chavez A, Dietz SL, Esvelt KM, Church GM (2015) Safeguarding CRISPR-Cas9 gene drives in yeast. *Nat Biotechnol* 33(12):1250–1255.

228. Gorter de Vries AR, et al. (2019) Allele-specific genome editing using CRISPR–Cas9 is associated with loss of heterozygosity in diploid yeast. *Nucleic Acids Res* 47(3):1362–1372.